GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 2, 2005, 08:41:35 ; Search time 171 Seconds
(without alignments)
1196.469 Million cell updates/sec

Title: US-10-677-865-1
Perfect score: 2752
Sequence: 1 MGSSRAPWMGRVGGHGMMAL.....GPQLADPEVFLPQGELLDSP 529

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2752	100.0	529	2	AAW80498	Aaw80498 A protein
2	2752	100.0	529	2	AAW85055	Aaw85055 Human pro
3	2752	100.0	529	4	AAU27550	Aau27550 Corneodes
4	2752	100.0	529	7	ABJ39148	Abj39148 Human cor
5	2752	100.0	529	8	ADN04226	Adn04226 Antipsori
6	2749	99.9	529	8	ADP12525	Adp12525 Protein e
7	2734	99.3	529	3	AAY44989	Aay44989 Human epi
8	2697	98.0	529	4	AAG64152	Aag64152 Polypepti
9	2505	91.0	486	7	ABJ39147	Abj39147 Human S p
10	2505	91.0	486	8	ADQ21504	Adq21504 Human sof
11	1106	40.2	219	8	ABO60412	Abo60412 Human gen
12	717	26.1	146	7	ABJ39195	Abj39195 Fragment
13	533	19.4	111	7	ABJ39196	Abj39196 Fragment
14	400	14.5	76	7	ABJ39197	Abj39197 Fragment
15	369	13.4	71	5	AAU91255	Aau91255 Nucleic a
16	340.5	12.4	950	4	ABB71513	Abb71513 Drosophil
17	320	11.6	2112	4	ABB60403	Abb60403 Drosophil
18	310	11.3	58	7	ABJ39194	Abj39194 Fragment
19	298	10.8	55	7	ABJ39193	Abj39193 Fragment
20	295	10.7	2280	4	ABB61650	Abb61650 Drosophil
21	289.5	10.5	1063	5	ABB54168	Abb54168 Lactococc
22	266	9.7	258	7	ADB64880	Adb64880 Human pro
23	265	9.6	400	6	AAO16495	Aao16495 Kukulcani
24	259	9.4	1010	2	AAW34989	Aaw34989 Teredinib
25	257	9.3	1459	8	ADN23765	Adn23765 Bacterial
26	257	9.3	2586	4	ABB66878	Abb66878 Drosophil
27	253.5	9.2	1217	6	ABP60377	Abp60377 Bombyx mo
28	249	9.0	1357	4	ABB70775	Abb70775 Drosophil
29	248.5	9.0	3507	8	ADN23131	Adn23131 Bacterial
30	248	9.0	627	7	ADB61322	Adb61322 Spider si
31	248	9.0	627	7	ADM46150	Adm46150 Nephila c
32	248	9.0	657	4	ABB57783	Abb57783 Drosophil
33	246	8.9	595	2	AAR14309	Aar14309 N.clavipe
34	246	8.9	595	2	AAW53347	Aaw53347 Nephila c
35	246	8.9	595	3	AAY59071	Aay59071 N. clavip
36	244.5	8.9	1253	3	AAB19772	Aab19772 Human den
37	244.5	8.9	1253	5	ABP51785	Abp51785 Human den
38	244	8.9	624	6	AAO16488	Aao16488 Nephila c
39	244	8.9	627	7	ADC35241	Adc35241 MaSpII si
40	244	8.9	1455	4	ABB63040	Abb63040 Drosophil
41	243.5	8.8	584	6	ABU22680	Abu22680 Protein e
42	243.5	8.8	1566	2	AAY06309	Aay06309 Human p53
43	243.5	8.8	1566	6	ABR47564	Abr47564 Breast ca
44	243.5	8.8	1566	8	ADH13181	Adh13181 Human mal
45	243.5	8.8	1566	8	ADO61338	Ado61338 Human PPA

ALIGNMENTS

RESULT 1
AAW80498
ID AAW80498 standard; protein; 529 AA.
XX
AC AAW80498;
XX
DT 10-FEB-1999 (first entry)
XX
DE A protein designated corneodesmosin.
XX
KW Corneodesmosin; intercorneocyte cohesion; skin trophic disorder;
KW scar formation; epidermis thinning; skin fragility; hyperkeratosis;
KW xerosis; ichthyosis; psoriasis; hyperkeratotic tumour;
KW reactive keratosis; and leucokeratosis.
XX
OS Homo sapiens.
XX
PN FR2761362-A1.
XX
PD 02-OCT-1998.
XX
PF 28-MAR-1997; 97FR-00003899.
XX
PR 28-MAR-1997; 97FR-00003899.
XX
PA (OREA) L'OREAL SA.
XX
PI Serre G, Simon M, Weber VM;
XX
DR WPI; 1998-586093/50.
DR N-PSDB; AAV63614.
XX
PT New polypeptide, corneodesmosin, and related nucleic acid - for cosmetic
PT or therapeutic treatment of trophic disorders or thinning of the
PT epidermis.
XX
PS Claim 1; Page 29; 35pp; French.
XX
CC The present sequence represents a protein designated corneodesmosin. The
CC protein is involved in intercorneocyte cohesion and is used, in cosmetic
CC or pharmaceutical compositions, to treat trophic disorders of the skin
CC and disorders associated with scar formation, and thinning of the
CC epidermis, particularly the corneal layer, or skin fragility, to
CC reinforce intercorneocyte cohesion and/or to induce thickening of the
CC corneal layer. Hyperkeratosis, xerosis, ichthyosis, psoriasis, malignant
CC or benign hyperkeratotic tumours (including those of the Malphigian
CC mucosa), reactive keratosis, and leucokeratosis (of the cervix uteri,
CC mouth or during prolapse) can also be treated with the products

XX
SQ Sequence 529 AA;

Query Match 100.0%; Score 2752; DB 2; Length 529;
Best Local Similarity 100.0%; Pred. No. 2.4e-165;
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60

Qy         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180

Qy        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240

Qy        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300

Qy        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360

Qy        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420

Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480

Qy        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
```

RESULT 2
AAW85055
ID AAW85055 standard; protein; 529 AA.
XX
AC AAW85055;
XX
DT 05-FEB-1999 (first entry)
XX
DE Human protein designated corneodesmosin.
XX
KW Corneodesmosin; intercorneocyte cohesion; skin trophic disorder;
KW scar formation; epidermis thinning; skin fragility; hyperkeratosis;
KW xerosis; ichthyosis; psoriasis; hyperkeratotic tumour;
KW reactive keratosis; and leucokeratosis.
XX
OS Homo sapiens.
XX
PN FR2761363-A1.
XX
PD 02-OCT-1998.
XX
PF 11-SEP-1997; 97FR-00011317.
XX
PR 28-MAR-1997; 97FR-00003899.
XX
PA (OREA) L'OREAL SA.
XX
PI Serre G, Simon M, Weber VM;
XX
DR WPI; 1998-586094/50.
DR N-PSDB; AAV63541.
XX
PT New polypeptide, corneodesmosin, from skin and related nucleic acid - for
PT cosmetic or therapeutic treatment of trophic disorders, thinning of the
PT epidermis etc., also compositions containing protease that degrades
PT corneodesmosin.
XX
PS Claim 1; Page 30; 35pp; French.
XX
CC The present sequence represents a protein designated corneodesmosin. The
CC protein is involved in intercorneocyte cohesion and is used, in cosmetic
CC or pharmaceutical compositions, to treat trophic disorders of the skin
CC and disorders associated with scar formation, and thinning of the
CC epidermis, particularly the corneal layer, or skin fragility, to
CC reinforce intercorneocyte cohesion and/or to induce thickening of the
CC corneal layer. Hyperkeratosis, xerosis, ichthyosis, psoriasis, malignant
CC or benign hyperkeratotic tumours (including those of the Malphigian
CC mucosa), reactive keratosis, and leucokeratosis (of the cervix uteri,
CC mouth or during prolapse) can also be treated with the products
XX
SQ Sequence 529 AA;

Query Match 100.0%; Score 2752; DB 2; Length 529;
Best Local Similarity 100.0%; Pred. No. 2.4e-165;
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60

Qy         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180

Qy        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240

Qy        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300

Qy        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360

Qy        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420

Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480

Qy        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
```

RESULT 3
AAU27550
ID AAU27550 standard; protein; 529 AA.
XX
AC AAU27550;
XX
DT 18-DEC-2001 (first entry)
XX
DE Corneodesmosin.

XX
KW Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW antipsoriatic; corneodesmosin; inflammatory disease; psoriasis.
XX
OS Homo sapiens.
XX
PN WO200162788-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000795.
XX
PR 23-FEB-2000; 2000GB-00004312.
XX
PA (OXAG-) OXAGEN LTD.
XX
PI Olaveson M, Lench N, Allen M, Tazi-Ahnini R;
XX
DR WPI; 2001-570627/64.
DR N-PSDB; AAS43749.
XX
PT Corneodesmosin protein and polynucleotide encoding it, having one or more
PT polymorphisms useful in treating, diagnosing or determining
PT susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT diseases.
XX
PS Claim 8; Fig 2; 60pp; English.
XX
CC The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC encoding the corneodesmosin gene, where the gene comprises a base
CC substitution, deletion or insertion at one or more positions. (I) and
CC (II) are useful for screening for agents for use in prognosis, diagnosis
CC and treatment of individuals having or being susceptible to
CC corneodesmosin-mediated disease, by monitoring the reaction between the
CC molecules and the agents. The nucleotide and amino acid polymorphisms are
CC useful for diagnosing or determining susceptible to corneodesmosin-
CC mediated disease, which facilitates subsequent treatment of the disease
CC for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC are useful in diagnostic, prognostic or therapeutic methods and as
CC research tools for e.g. in drug screening. (II) is useful as probes or
CC primers for detecting an allele of the polymorphism or in the regulation
CC of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC screening DNA clone libraries for cells secreting the antigen. (II) is
CC useful as a model to investigate the role of corneodesmosin in normal
CC skin function. The present sequence represents corneodesmosin amino acid
CC sequence, as described in the invention
XX
SQ Sequence 529 AA;

Query Match 100.0%; Score 2752; DB 4; Length 529;
Best Local Similarity 100.0%; Pred. No. 2.4e-165;
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60

Qy         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180

Qy        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240

Qy        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300

Qy        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360

Qy        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420

Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480

Qy        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
```

RESULT 4
ABJ39148
ID ABJ39148 standard; protein; 529 AA.
XX
AC ABJ39148;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human corneodesmosin (CDSN) protein.
XX
KW Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW winter xerosis; ichthyoses; dermatological; gene therapy; chromosome 6.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_difference 28
FT /note= "Encoded by CCAG"
FT misc_difference 408
FT /note= "Encoded by TCC"
XX
PN WO2003050311-A2.
XX
PD 19-JUN-2003.
XX
PF 09-DEC-2002; 2002WO-US039485.
XX
PR 07-DEC-2001; 2001US-0338188P.
XX
PA (WELA) WELLA AG.
XX
PI Pras E, Levy-Nissenbaum E, Frydman M, Thederahn T;
XX
DR WPI; 2003-569100/53.
DR N-PSDB; ABT43849, ABT43850.
XX
PT Therapeutic composition comprising an effector of a corneodesmosin gene
PT function associated with hair growth and a carrier, useful for treating
PT alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS Disclosure; Fig 2e; 37pp; English.
XX
CC This invention relates to the identification of genes linked to hair loss
CC and the retardation of hair growth. Alopecia is the general term used to
CC refer to loss of hair, the causes of which can include autoimmune
CC disorders, hormonal imbalances, neurological conditions and genetic
CC abnormalities. Specifically, the invention describes the highly conserved
CC corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC in the onset of nonsyndromic alopecia, an example being hytrichosis
CC simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC the dermatological activity of CDSN provides a good diagnostic and
CC therapeutic target for alopecia, and through gene therapy CDSN can be

```
CC   used to modulate hair growth and treat male pattern baldness. This
CC   polypeptide sequence is the human corneodesmosin (CDSN) gene of the
CC   invention
XX
SQ   Sequence 529 AA;

  Query Match            100.0%;  Score 2752;  DB 7;  Length 529;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-165;
  Matches  529;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60

Qy         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180

Qy        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240

Qy        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300

Qy        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360

Qy        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420

Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480

Qy        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 5
ADN04226
ID   ADN04226 standard; protein; 529 AA.
XX
AC   ADN04226;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Antipsoriatic protein sequence #308.
XX
KW   antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2004028479-A2.
XX
PD   08-APR-2004.
XX
PF   25-SEP-2003; 2003WO-US030907.
XX
PR   25-SEP-2002; 2002US-0414006P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Bodary S,  Clark H,  Jackman J,  Schoenfeld J,  Williams PM,  Wood WI;
PI   Wu TD;
XX
DR   WPI; 2004-305105/28.
DR   N-PSDB; ADN04225.
XX
PT   New PRO nucleic acid or polypeptide, useful for preparing a
PT   pharmaceutical composition for diagnosing or treating psoriasis in a
PT   mammal.
XX
PS   Claim 9; SEQ ID NO 620; 3069pp; English.
XX
CC   The invention relates to novel polynucleotide and polypeptides for
CC   treating psoriasis or a sequence having at least 80% identity to the
CC   above sequences. The nucleic acid is useful for preparing a composition
CC   for diagnosing or treating psoriasis in a mammal. This sequence
CC   corresponds to one of the polypeptides of the invention.
XX
SQ   Sequence 529 AA;

  Query Match            100.0%;  Score 2752;  DB 8;  Length 529;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-165;
  Matches  529;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60

Qy         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180

Qy        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240

Qy        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300

Qy        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360

Qy        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420

Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480

Qy        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 6
ADP12525
ID   ADP12525 standard; protein; 529 AA.
XX
AC   ADP12525;
XX
DT   12-AUG-2004  (first entry)
XX
DE   Protein encoded by mRNA of the invention #135.
XX
KW   transplant rejection; immune system; rheumatoid arthritis; lupus;
KW   inflammatory bowel disease; multiple sclerosis; HIV; AIDS.
XX
OS   Homo sapiens.
```

XX
PN WO2004042346-A2.
XX
PD 21-MAY-2004.
XX
PF 24-APR-2003; 2003WO-US012946.
XX
PR 24-APR-2002; 2002US-00131831.
PR 20-DEC-2002; 2002US-00325899.
XX
PA (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
PI Rosenberg S;
XX
DR WPI; 2004-400724/37.
XX
PT Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT rejection, in an individual, comprises detecting the expression level of
PT the genes.
XX
PS Claim 65; SEQ ID NO 2534; 1762pp; English.
XX
CC The present invention relates to diagnosing or monitoring transplant
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC comprises detecting the expression level of one or more genes. The
CC methods, system and kits are useful in diagnosing or monitoring
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC islet, lung, bone marrow or stem cell transplant rejection,
CC xenotransplant rejection or mechanical organ replacement rejection, in an
CC individual. The method is also useful in assessing the immune status of
CC an individual. The methods are also useful in diagnosing and monitoring
CC diseases that involve the immune system, e.g. rheumatoid arthritis,
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC viral, bacterial or fungal infection. The present sequence represents a
CC protein that is encoded by the mRNA of the invention.
XX
SQ Sequence 529 AA;

Query Match 99.9%; Score 2749; DB 8; Length 529;
Best Local Similarity 99.8%; Pred. No. 3.7e-165;
Matches 528; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60

Qy         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180

Qy        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240

Qy        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300

Qy        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360

Qy        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGSPYHPCGSASQ 420

Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480

Qy        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
```

RESULT 7
AAY44989
ID AAY44989 standard; protein; 529 AA.
XX
AC AAY44989;
XX
DT 23-MAY-2000 (first entry)
XX
DE Human epidermal protein-6.
XX
KW Human epidermal protein-6; HEPI; epithelial disorder; scabies;
KW dyshidrotic eczema; cell proliferative disorder; actinic keratosis;
KW arteriosclerosis; autoimmune disorder; inflammatory disorder;
KW acquired immune deficiency syndrome; AIDS; Addison's disease; antiHIV;
KW dermatological; anitarteriosclerotic; antiinflammatory;
KW immunosuppressive.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	Peptide	1. .32
FT		/label= Signal_peptide
FT	Modified-site	3
FT		/note= "Potential phosphorylation site"
FT	Modified-site	37
FT		/note= "Potential phosphorylation site"
FT	Modified-site	50
FT		/note= "Potential phosphorylation site"
FT	Modified-site	57
FT		/note= "Potential phosphorylation site"
FT	Modified-site	80
FT		/note= "Potential phosphorylation site"
FT	Modified-site	105
FT		/note= "Potential phosphorylation site"
FT	Modified-site	172
FT		/note= "N-glycosylated"
FT	Modified-site	182
FT		/note= "Potential phosphorylation site"
FT	Modified-site	287
FT		/note= "Potential phosphorylation site"
FT	Modified-site	395
FT		/note= "Potential phosphorylation site"
FT	Modified-site	440
FT		/note= "Potential phosphorylation site"
FT	Modified-site	498
FT		/note= "Potential phosphorylation site"
FT	Modified-site	498
FT		/note= "Potential phosphorylation site"

XX
PN WO200006727-A2.
XX
PD 10-FEB-2000.
XX
PF 27-JUL-1999; 99WO-US017107.
XX
PR 28-JUL-1998; 98US-0155203P.
PR 07-DEC-1998; 98US-0155254P.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Tang YT, Lal P, Corley NC, Guegler KJ, Patterson C, Baughn MR;
PI Yue H;
XX
DR WPI; 2000-195295/17.
DR N-PSDB; AAZ50582.

```
XX
PT   New human epidermal proteins (HEPI-1) to (HEPI-6) useful for the
PT   diagnosis, treatment and prevention of epithelial, cell proliferative,
PT   and autoimmune inflammatory disorders.
XX
PS   Claim 1; Fig 4; 82pp; English.
XX
CC   The present sequence is human epidermal protein-6 (HEPI) expressed in
CC   breast skin and keratinocytes. This protein is derived from PTHYNOT03
CC   cDNA library. HEPI-6 shares 99% identity with human S protein.
CC   Recombinant vectors comprising HEPI cDNA are introduced into host cells
CC   for protein expression. The HEPI proteins are useful for the treatment of
CC   epithelial disorders, including dyshidrotic eczema and scabies, cell
CC   proliferative disorders including actinic keratosis and arteriosclerosis,
CC   and autoimmune/inflammatory disorders like acquired immune deficiency
CC   syndrome (AIDS) and Addison's disease. Pharmaceutical compositions
CC   comprising HEPI proteins are useful for treating disorders associated
CC   with altered HEPI expression
XX
SQ   Sequence 529 AA;

  Query Match             99.3%;  Score 2734;  DB 3;  Length 529;
  Best Local Similarity   99.2%;  Pred. No. 3.2e-164;
  Matches  525;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db          1 MGSSRAPWMGRVGGHGMLALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60

Qy         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180

Qy        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        181 NSYRGILNPSQPGQSSSSSQTFGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240

Qy        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300

Qy        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360

Qy        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
              ||||||||||||||||||||||||||||||||||||||||||||:| ||||||||||
Db        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGLPYHPCGSASQ 420

Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480

Qy        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 8
AAG64152
ID   AAG64152 standard; protein; 529 AA.
XX
AC   AAG64152;
XX
DT   01-OCT-2001  (first entry)
XX
DE   Polypeptide #1 for vulgar psoriasis diagnosis.
XX
KW   MHC S; major histocompatibility complex S; vulgar psoriasis; diagnosis;
KW   primer; SEEK1; HCR; a-helix coiled-coil rod homologue; polymorphism.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 143
FT                   /label= Ser, Asn
FT   Misc-difference 151
FT                   /label= Ser, Cys
FT   Misc-difference 152
FT                   /label= Ser, Arg
FT   Misc-difference 200
FT                   /label= Ser, Phe
FT   Misc-difference 282
FT                   /label= Pro, Thr
FT   Misc-difference 401
FT                   /label= Ser, Gly
FT   Misc-difference 408
FT                   /label= Ala, Ser
FT   Misc-difference 409
FT                   /label= Gly, Val
FT   Misc-difference 410
FT                   /label= Ser, Leu
FT   Misc-difference 453
FT                   /label= Ser, Asn
FT   Misc-difference 527
FT                   /label= Asp, Asn
XX
PN   WO200142458-A1.
XX
PD   14-JUN-2001.
XX
PF   06-DEC-2000; 2000WO-JP008624.
XX
PR   06-DEC-1999;   99JP-00346867.
XX
PA   (INOK/) INOKO H.
XX
PI   Inoko H,  Tamiya G;
XX
DR   WPI; 2001-381680/40.
DR   N-PSDB; AAH73624.
XX
PT   New primer DNA, useful for detecting vulgar psoriasis.
XX
PS   Disclosure; Fig 2; 106pp; Japanese.
XX
CC   The invention relates to a method of diagnosing vulgar psoriasis using
CC   primers based on the sequences of the human MHC S, SEEK1 and HCR genes.
CC   By analysing the sequences of these genes in Japanese patients with
CC   psoriasis and in normal subjects, it has been found that some of the
CC   examined polymorphisms correlate significantly to the group of patients
CC   with psoriasis. Vulgar psoriasis can therefore be diagnosed by analysing
CC   these gene polymorphisms. The present sequence is provided in the
CC   specification
XX
SQ   Sequence 529 AA;

  Query Match             98.0%;  Score 2697;  DB 4;  Length 529;
  Best Local Similarity   97.9%;  Pred. No. 7e-162;
  Matches  518;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60

Qy         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
              |||||||||||||||||||| ||||||| ||||||||||||||||||||||||||||||
Db        121 SGSSLQGASGSSQLGSSSSHSGXSGSHSGSXXSHSSSSSSFQFSSSSFQVGNGSALPTND 180
```

```
Qy        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        181 NSYRGILNPSQPGQSSSSSQTXGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240

Qy        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKXCPPITSVDKSYGGYEVVG 300

Qy        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360

Qy        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
              |||||||||||||||||||||||||||||||||||||| ||||||   ||||||||||
Db        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSXSSISSSXXXPYHPCGSASQ 420

Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSXSSGHPCMSVSSLTLTGGPDGSPHPDPS 480

Qy        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
              ||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLXSP 529
```

RESULT 9
ABJ39147
ID ABJ39147 standard; protein; 486 AA.
XX
AC ABJ39147;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human S protein encoded by the skin corneodesmosin (CDSN) gene (S-gene).
XX
KW Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW winter xerosis; ichthyoses; dermatological; gene therapy; chromosome 6;
KW S protein; skin.
XX
OS Homo sapiens.
XX
PN WO2003050311-A2.
XX
PD 19-JUN-2003.
XX
PF 09-DEC-2002; 2002WO-US039485.
XX
PR 07-DEC-2001; 2001US-0338188P.
XX
PA (WELA) WELLA AG.
XX
PI Pras E, Levy-Nissenbaum E, Frydman M, Thederahn T;
XX
DR WPI; 2003-569100/53.
DR N-PSDB; ABT43848.
XX
PT Therapeutic composition comprising an effector of a corneodesmosin gene
PT function associated with hair growth and a carrier, useful for treating
PT alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS Disclosure; Fig 2b; 37pp; English.
XX
CC This invention relates to the identification of genes linked to hair loss
CC and the retardation of hair growth. Alopecia is the general term used to
CC refer to loss of hair, the causes of which can include autoimmune
CC disorders, hormonal imbalances, neurological conditions and genetic
CC abnormalities. Specifically, the invention describes the highly conserved
CC corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC in the onset of nonsyndromic alopecia, an example being hytrichosis
CC simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC the dermatological activity of CDSN provides a good diagnostic and
CC therapeutic target for alopecia, and through gene therapy CDSN can be
CC used to modulate hair growth and treat male pattern baldness. This
CC polypeptide sequence is the human S protein (corneodesmosin protein only
CC expressed in the skin), the target of the invention
XX
SQ Sequence 486 AA;

Query Match 91.0%; Score 2505; DB 7; Length 486;
Best Local Similarity 99.6%; Pred. No. 8.2e-150;
Matches 482; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy         17 MMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGS 76
              |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGS 60

Qy         77 SISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGS 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGS 120

Qy        137 SSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSS 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSS 180

Qy        197 SSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVV 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVV 240

Qy        257 VDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKG 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKG 300

Qy        317 KIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGG 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGG 360

Qy        377 VQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSS 436
              |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        361 VQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGSPYHPCGSASQSPCSPPGTGSFSSSSS 420

Qy        437 SQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPC 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPC 480

Qy        497 RSIR 500
              ||||
Db        481 RSIR 484
```

RESULT 10
ADQ21504
ID ADQ21504 standard; protein; 486 AA.
XX
AC ADQ21504;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated protein - SEQ ID 4324.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX
OS Homo sapiens.
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX

PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 4324; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC protein of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 486 AA;

Query Match 91.0%; Score 2505; DB 8; Length 486;
Best Local Similarity 99.6%; Pred. No. 8.2e-150;
Matches 482; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy         17 MMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGS 76
              |:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGS 60

Qy         77 SISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGS 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGS 120

Qy        137 SSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSS 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSS 180

Qy        197 SSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVV 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVV 240

Qy        257 VDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKG 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKG 300

Qy        317 KIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGG 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGG 360

Qy        377 VQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSS 436
              ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        361 VQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGSPYHPCGSASQSPCSPPGTGSFSSSSS 420

Qy        437 SQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPC 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPC 480

Qy        497 RSIR 500
              ||||
Db        481 RSIR 484
```

RESULT 11
ABO60412
ID ABO60412 standard; protein; 219 AA.
XX
AC ABO60412;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon protein #6646.
XX
KW Human; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 45; SEQ ID NO 34046; 80pp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above- mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe protein of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 219 AA;

```
  Query Match             40.2%;  Score 1106;  DB 8;  Length 219;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-62;
  Matches  219;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         30 TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 60

Qy         90 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG 149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG 120

Qy        150 SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQ 209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQ 180

Qy        210 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSG 248
              ||||||||||||||||||||||||||||||||||||||||
Db        181 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSG 219


RESULT 12
ABJ39195
ID   ABJ39195 standard; protein; 146 AA.
XX
AC   ABJ39195;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Fragment c, cleavage product of human S (corneodesmosin skin) protein.
XX
KW   Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW   neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW   hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW   winter xerosis; ichthyoses; dermatological; gene therapy; S protein;
KW   skin.
XX
OS   Homo sapiens.
XX
PN   WO2003050311-A2.
XX
PD   19-JUN-2003.
XX
PF   09-DEC-2002; 2002WO-US039485.
XX
PR   07-DEC-2001; 2001US-0338188P.
XX
PA   (WELA ) WELLA AG.
XX
PI   Pras E,  Levy-Nissenbaum E,  Frydman M,  Thederahn T;
XX
DR   WPI; 2003-569100/53.
XX
PT   Therapeutic composition comprising an effector of a corneodesmosin gene
PT   function associated with hair growth and a carrier, useful for treating
PT   alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS   Disclosure; Fig 2e; 37pp; English.
XX
CC   This invention relates to the identification of genes linked to hair loss
CC   and the retardation of hair growth. Alopecia is the general term used to
CC   refer to loss of hair, the causes of which can include autoimmune
CC   disorders, hormonal imbalances, neurological conditions and genetic
CC   abnormalities. Specifically, the invention describes the highly conserved
CC   corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC   in the onset of nonsyndromic alopecia, an example being hytrichosis
CC   simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC   chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC   including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC   the dermatological activity of CDSN provides a good diagnostic and
CC   therapeutic target for alopecia, and through gene therapy CDSN can be
CC   used to modulate hair growth and treat male pattern baldness. This
CC   peptide sequence is a functional region of the human S protein, the
CC   cornesodesmosin protein solely expressed in the skin that is a target of
CC   the invention
XX
SQ   Sequence 146 AA;

  Query Match             26.1%;  Score 717;  DB 7;  Length 146;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-37;
  Matches  146;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         55 CLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQ 114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQ 60

Qy        115 VSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGS 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGS 120

Qy        175 ALPTNDNSYRGILNPSQPGQSSSSSQ 200
              ||||||||||||||||||||||||||
Db        121 ALPTNDNSYRGILNPSQPGQSSSSSQ 146


RESULT 13
ABJ39196
ID   ABJ39196 standard; protein; 111 AA.
XX
AC   ABJ39196;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Fragment d, glycine rich sticky domain 1 of the human S protein.
XX
KW   Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW   neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW   hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW   winter xerosis; ichthyoses; dermatological; gene therapy; S protein;
KW   skin.
XX
OS   Homo sapiens.
XX
PN   WO2003050311-A2.
XX
PD   19-JUN-2003.
XX
PF   09-DEC-2002; 2002WO-US039485.
XX
PR   07-DEC-2001; 2001US-0338188P.
XX
PA   (WELA ) WELLA AG.
XX
PI   Pras E,  Levy-Nissenbaum E,  Frydman M,  Thederahn T;
XX
DR   WPI; 2003-569100/53.
XX
PT   Therapeutic composition comprising an effector of a corneodesmosin gene
PT   function associated with hair growth and a carrier, useful for treating
PT   alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS   Disclosure; Fig 2e; 37pp; English.
XX
CC   This invention relates to the identification of genes linked to hair loss
CC   and the retardation of hair growth. Alopecia is the general term used to
CC   refer to loss of hair, the causes of which can include autoimmune
CC   disorders, hormonal imbalances, neurological conditions and genetic
CC   abnormalities. Specifically, the invention describes the highly conserved
CC   corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC   in the onset of nonsyndromic alopecia, an example being hytrichosis
CC   simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC   chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC   including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC   the dermatological activity of CDSN provides a good diagnostic and
CC   therapeutic target for alopecia, and through gene therapy CDSN can be
CC   used to modulate hair growth and treat male pattern baldness. This
```

```
CC   peptide sequence is a functional region of the human S protein, the
CC   cornesodesmosin protein solely expressed in the skin that is a target of
CC   the invention
XX
SQ   Sequence 111 AA;

  Query Match            19.4%;  Score 533;  DB 7;  Length 111;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-26;
  Matches  111;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         65 FSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSS 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 FSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSS 60

Qy        125 LQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSA 175
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSA 111


RESULT 14
ABJ39197
ID   ABJ39197 standard; protein; 76 AA.
XX
AC   ABJ39197;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Fragment e, glycine rich sticky domain 2 of the human S protein.
XX
KW   Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW   neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW   hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW   winter xerosis; ichthyoses; dermatological; gene therapy; S protein;
KW   skin.
XX
OS   Homo sapiens.
XX
PN   WO2003050311-A2.
XX
PD   19-JUN-2003.
XX
PF   09-DEC-2002; 2002WO-US039485.
XX
PR   07-DEC-2001; 2001US-0338188P.
XX
PA   (WELA ) WELLA AG.
XX
PI   Pras E,  Levy-Nissenbaum E,  Frydman M,  Thederahn T;
XX
DR   WPI; 2003-569100/53.
XX
PT   Therapeutic composition comprising an effector of a corneodesmosin gene
PT   function associated with hair growth and a carrier, useful for treating
PT   alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS   Disclosure; Fig 2e; 37pp; English.
XX
CC   This invention relates to the identification of genes linked to hair loss
CC   and the retardation of hair growth. Alopecia is the general term used to
CC   refer to loss of hair, the causes of which can include autoimmune
CC   disorders, hormonal imbalances, neurological conditions and genetic
CC   abnormalities. Specifically, the invention describes the highly conserved
CC   corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC   in the onset of nonsyndromic alopecia, an example being hytrichosis
CC   simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC   chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC   including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC   the dermatological activity of CDSN provides a good diagnostic and
CC   therapeutic target for alopecia, and through gene therapy CDSN can be
CC   used to modulate hair growth and treat male pattern baldness. This
CC   peptide sequence is a functional region of the human S protein, the
CC   cornesodesmosin protein solely expressed in the skin that is a target of
CC   the invention
XX
SQ   Sequence 76 AA;

  Query Match            14.5%;  Score 400;  DB 7;  Length 76;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-18;
  Matches   76;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        375 GGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSS 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSS 60

Qy        435 SSSQSSGKIILQPCGS 450
              ||||||||||||||||
Db         61 SSSQSSGKIILQPCGS 76


RESULT 15
AAU91255
ID   AAU91255 standard; protein; 71 AA.
XX
AC   AAU91255;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Nucleic acid selection related corneodesmosin peptide.
XX
KW   Nucleic acid selection; viral genome; sindbis virus.
XX
OS   Mus musculus.
XX
PN   WO200216572-A2.
XX
PD   28-FEB-2002.
XX
PF   17-AUG-2001; 2001WO-CA001169.
XX
PR   18-AUG-2000; 2000US-00641931.
XX
PA   (PHEN-) PHENOGENE THERAPEUTIQUES INC.
XX
PI   Lanctot C,  Moffat P,  Salois P;
XX
DR   WPI; 2002-257913/30.
XX
PT   Selecting exogenous nucleic acids having a desired feature, particularly
PT   from a library of nucleic acids, involves using a viral genome.
XX
PS   Example 4; Page 55; 88pp; English.
XX
CC   The invention describes the use of a viral genome for selecting an
CC   exogenous nucleic acid having a desired feature. The method requires a
CC   viral genome, e.g. sindbis virus, capable of expressing the exogenous
CC   nucleic acid in a host cell and selecting among these the ones having the
CC   desired feature or predefined function. The method removes the expensive
CC   and time-consuming task of selecting cells that express a gene of
CC   interest. The method is more rapid, efficient and accurate for selecting
CC   a particular nucleic acid having a desired feature, characteristic or
CC   function. This sequence represents a peptide encoded by a nucleic acid
CC   isolated using the method described in the invention
XX
SQ   Sequence 71 AA;

  Query Match            13.4%;  Score 369;  DB 5;  Length 71;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-16;
  Matches   69;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60

Qy         61 DSSGFSSYS 69
              |||||||||
Db         61 DSSGFSSYS 69
```

Search completed: August 2, 2005, 09:09:22
Job time : 176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 2, 2005, 09:01:18 ; Search time 44 Seconds
(without alignments)
897.485 Million cell updates/sec

Title: US-10-677-865-1
Perfect score: 2752
Sequence: 1 MGSSRAPWMGRVGGHGMMAL.....GPQLADPEVFLPQGELLDSP 529

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2752	100.0	529	4	US-09-381-656-1	Sequence 1, Appli
2	282.5	10.3	78	4	US-09-641-931F-39	Sequence 39, Appl
3	248.5	9.0	493	3	US-08-556-978B-59	Sequence 59, Appl
4	246	8.9	595	1	US-08-425-069-4	Sequence 4, Appli
5	246	8.9	595	2	US-08-317-844B-4	Sequence 4, Appli
6	244.5	8.9	529	3	US-09-247-806-2	Sequence 2, Appli
7	243.5	8.8	1566	4	US-09-581-472B-2	Sequence 2, Appli
8	243.5	8.8	1581	3	US-09-110-517-2	Sequence 2, Appli
9	243	8.8	1259	4	US-09-949-016-10366	Sequence 10366, A
10	241.5	8.8	316	4	US-09-538-092-997	Sequence 997, App
11	239.5	8.7	461	4	US-09-248-796A-23039	Sequence 23039, A
12	225.5	8.2	100	4	US-09-411-067C-5	Sequence 5, Appli
13	223.5	8.1	1306	4	US-09-538-092-330	Sequence 330, App
14	221	8.0	1970	4	US-09-538-092-1005	Sequence 1005, Ap
15	219.5	8.0	1475	4	US-09-538-092-1160	Sequence 1160, Ap
16	218	7.9	1011	1	US-08-477-509B-94	Sequence 94, Appl
17	218	7.9	1011	3	US-08-482-085B-94	Sequence 94, Appl
18	218	7.9	1011	3	US-09-444-791A-94	Sequence 94, Appl
19	218	7.9	1170	1	US-08-175-155-59	Sequence 59, Appl
20	218	7.9	1170	2	US-08-707-237A-66	Sequence 66, Appl
21	216.5	7.9	1060	3	US-08-911-393-2	Sequence 2, Appli
22	216.5	7.9	1060	4	US-09-955-909-2	Sequence 2, Appli
23	215	7.8	2090	4	US-09-538-092-1081	Sequence 1081, Ap
24	215	7.8	2120	4	US-09-949-016-9768	Sequence 9768, Ap
25	213	7.7	552	3	US-09-219-849-7	Sequence 7, Appli
26	211.5	7.7	960	3	US-09-219-849-6	Sequence 6, Appli
27	211	7.7	870	2	US-09-010-928B-2	Sequence 2, Appli
28	209	7.6	1343	4	US-09-949-016-10641	Sequence 10641, A
29	209	7.6	2018	3	US-09-444-791A-80	Sequence 80, Appl
30	209	7.6	2100	1	US-08-477-509B-80	Sequence 80, Appl
31	209	7.6	2100	3	US-08-482-085B-80	Sequence 80, Appl
32	209	7.6	2107	1	US-08-175-155-45	Sequence 45, Appl
33	209	7.6	2107	2	US-08-707-237A-51	Sequence 51, Appl
34	208.5	7.6	682	1	US-08-642-255-126	Sequence 126, App
35	208.5	7.6	682	1	US-08-397-633A-36	Sequence 36, Appl
36	208	7.6	714	3	US-08-556-978B-61	Sequence 61, Appl
37	208	7.6	714	3	US-09-247-806-10	Sequence 10, Appl
38	207.5	7.5	907	2	US-09-010-928B-4	Sequence 4, Appli
39	207	7.5	633	1	US-08-642-255-73	Sequence 73, Appl
40	207	7.5	1169	3	US-08-806-029-33	Sequence 33, Appl
41	206	7.5	619	4	US-09-252-991A-26352	Sequence 26352, A
42	205.5	7.5	749	1	US-08-317-522A-2	Sequence 2, Appli
43	205.5	7.5	749	1	US-08-439-818A-2	Sequence 2, Appli
44	205.5	7.5	749	2	US-08-751-965-2	Sequence 2, Appli
45	205.5	7.5	749	2	US-08-738-975-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-381-656-1
; Sequence 1, Application US/09381656
; Patent No. 6645509
; GENERAL INFORMATION:
; APPLICANT: SERRE, Guy Bruno Rene
; APPLICANT: SIMON, Michel
; APPLICANT: WEBER-VIVAT, Marina
; TITLE OF INVENTION: POLYPEPTIDE EXPRESSED IN THE HORNY LAYER OF EPIDERMIS
; TITLE OF INVENTION: AND USE THEREOF
; FILE REFERENCE: 016800-336
; CURRENT APPLICATION NUMBER: US/09/381,656
; CURRENT FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: FR 97/03899
; PRIOR FILING DATE: 1997-03-28
; PRIOR APPLICATION NUMBER: FR 97/11317
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Human
US-09-381-656-1

Query Match 100.0%; Score 2752; DB 4; Length 529;
Best Local Similarity 100.0%; Pred. No. 6.6e-175;
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60

Qy         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180

Qy        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240

Qy        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300

Qy        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360

Qy        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420

Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480

Qy        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 2
US-09-641-931F-39
; Sequence 39, Application US/09641931F
; Patent No. 6830885
; GENERAL INFORMATION:
;  APPLICANT: Lanctot, et al.
;  TITLE OF INVENTION: Nucleic Acid Molecule, Method and Kit for Selecting a
;  TITLE OF INVENTION:  Nucleic Acid Having A Desired Feature
;  FILE REFERENCE: 2003390-0001
;  CURRENT APPLICATION NUMBER: US/09/641,931F
;  CURRENT FILING DATE:  2003-04-08
;  NUMBER OF SEQ ID NOS: 45
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
;   LENGTH: 78
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-641-931F-39

  Query Match             10.3%;  Score 282.5;  DB 4;  Length 78;
  Best Local Similarity   72.7%;  Pred. No. 4.8e-12;
  Matches   56;  Conservative    7;  Mismatches   13;  Indels    1;  Gaps    1;

Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              ||||:|| || |||||:||||:|| :||| ||||||| |||||||||||||||||| ||
Db          1 MGSSQAPRMGSVGGHGLMALLMAG-ILPGILAKSIGTLSDPCKDPTRITSPNDPCLIGKT 59

Qy         61 DSSGFSSYSGSSSSGSS 77
                |:  ||  ||||  ::
Db         60 GSNSISSQGGSSSRSAA 76


RESULT 3
US-08-556-978B-59
; Sequence 59, Application US/08556978B
; Patent No. 6268169
;  GENERAL INFORMATION:
;    APPLICANT:  FAHNESTOCK, STEPHEN F.
;    TITLE OF INVENTION:  NOVEL RECOMBINANTLY PRODUCED
;    TITLE OF INVENTION:  SPIDER SILK ANALOGS
;    NUMBER OF SEQUENCES:  107
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E. I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WINDOWS 95
;      SOFTWARE:  MICROSOFT WORD FOR WINDOWS 95
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/556,978B
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/077,600
;      FILING DATE:  JUNE 15, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FLOYD, LINDA AXAMETHY
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CR-9389-A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  59:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  493 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-556-978B-59

  Query Match             9.0%;  Score 248.5;  DB 3;  Length 493;
  Best Local Similarity   25.6%;  Pred. No. 7.5e-09;
  Matches  121;  Conservative   51;  Mismatches  216;  Indels   85;  Gaps   20;

Qy         58 GKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQG----------GSAGSF 106
               | |       |  ||::: |: :|| | | | |          |          | | |
Db         62 GYGQGQQGPSGPGSAAAASAAASAESGGPGGYGPGQQGPGGYGPGQQGPGGYGPGQQGPS 121

Qy        107 KPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSS 166
               ||:  :  : :|| |    | | | |    | |    |:: :::::|
Db        122 GPGSAAAAAAAAASGPGQQGPGGYGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAS------ 175

Qy        167 SFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSP 226
                  | |   |       |   | | |         | |  | : ::
Db        176 ----GPGQQGP-------GGYGPGQQGPGGYGPGQQGTSGPGSAAAA------------- 211

Qy        227 CSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSG--APGVVQGPPCSNGGLPGKPCP 284
                : ||      | | |     ||             | |   ||  |||     | |   |
Db        212 AAAGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAAGPGGYGPG-QQGPGGYGPGQQGPSGP 270

Qy        285 PITSVDKSYG-GYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGP 343
                :   : | | : :||       ||  |  |:  | ||     |  ||    : ||||
Db        271 GSAAAAAAAGPGQQGLGGYGPGQQGPG-GYGPGQQGPGGY----GP--GSASAAAAAAGP 323

Qy        344 ------PISEGKYFSSNPIIPSQSAASSAIAFQPVGTG-GVQLCGGGSTGSKGPCSPSSS 396
                     |:   | |   | |||::| |  | | | | |  || : | :||  | |
Db        324 GQQGPGGYGPGQQGPSGP--GSASAAAAAAAAGPGGYGPGQQGPGGYAPGQQGPSGPGS- 380

Qy        397 RVPSSSSISSSAGSP--YHP----CGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCGS 450
                 :|:: :::|  |  | |     |  :     | | ||  ::::::  |:|     | |
Db        381 ---ASAAAAAAAAGPGGYGPGQQGPGGYAPGQQGPSGPGS-AAAAAAASAG-----PGGY 431

Qy        451 KSSSSGHPCMSVSSLTLTGGPDG-SPHPDPSAGAKP-------CGSSSAGKIP 495
                  :  |     :::   : || |  |      ||  |          |: |||  |
Db        432 GPAQQGPSGPGIAASAASAGPGGYGPAQQGPAGYGPGSAVAASAGAGSAGYGP 484


RESULT 4
US-08-425-069-4
; Sequence 4, Application US/08425069
; Patent No. 5728810
;  GENERAL INFORMATION:
;    APPLICANT:  Lewis, Randolph V.
;    APPLICANT:  Xu, Ming
;    APPLICANT:  Hinman, Michael B.
;    TITLE OF INVENTION:  ISOLATED DNA CODING FOR SPIDER SILK
;    TITLE OF INVENTION:  PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
;    TITLE OF INVENTION:  CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolasch & Birch
;      STREET:  301 No. 5728810th Washington Street
;      CITY:  Falls Church
;      STATE:  Virginia
```

```
;      COUNTRY:  U.S.A.
;      ZIP:  22046
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/425,069
;      FILING DATE:  19-APR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murphy Jr., Gerald M
;      REGISTRATION NUMBER:  28,977
;      REFERENCE/DOCKET NUMBER:  1447-106P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 205-8000
;      TELEFAX:  (703) 205-8050
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  595 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-425-069-4

  Query Match             8.9%;  Score 246;  DB 1;  Length 595;
  Best Local Similarity   25.1%;  Pred. No. 1.4e-08;
  Matches  126;  Conservative   56;  Mismatches  217;  Indels  102;  Gaps   23;

Qy         58 GKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQG----------GSAG 104
              | |      |  ||::: |: :|| |   |  |   |      |           | |
Db         98 GYGQGQQGPSGPGSAAAASAAASAESGQQGPGGYGPGQQGPGGYGPGQQGPGGYGPGQQG 157

Qy        105 SFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFS 164
                ||:  :  : :|| |    |  |  | |       | |     |:: :::::|
Db        158 PSGPGSAAAAAAAASGPGQQGPGGYGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAS---- 213

Qy        165 SSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPD 224
                  | |   |       |   | | |         |:|  | : ::
Db        214 ------GPGQQGP-------GGYGPGQQGPGGYGPGQQGLSGPGSAAAA---------- 249

Qy        225 SPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSG--APGVVQGPPCSNGGL-PGK 281
                : ||    | | |    ||          | |   ||  |||    || ||:
Db        250 --AAAGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAAGPGGYGPG-QQGP----GGYGPGQ 302

Qy        282 PCP----PITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVP 337
              |        :   :  | : :||        ||  |  |:  | ||      |  ||
Db        303 QGPSGAGSAAAAAAAGPGQQGLGGYGPGQQGPG-GYGPGQQGPGGY----GP--GSASAA 355

Qy        338 SFAAGP------PISEGKYFSSNPIIPSQSAASSAIAFQPVGTG-GVQLCGGGSTGSKGP 390
              : ||||           |:   | |   | |||::| |  | | | | |  || : | :||
Db        356 AAAAGPGQQGPGGYGPGQQGPSGP--GSASAAAAAAAAGPGGYGPGQQGPGGYAPGQQGP 413

Qy        391 CSPSSSRVPSSSSISSSAGSP--YHP----CGSASQSPCSPPGTGSFSSSSSSQSSGKII 444
                | |   :|:: :::|  |  | |      |  :     | | || :::::: |:|
Db        414 SGPGS----ASAAAAAAAAGPGGYGPGQQGPGGYAPGQQGPSGPGS-AAAAAAASAG--- 465

Qy        445 LQPCGSKSSSSGHPCMSVSSLTLTGGPDG-SPHPDPSAGAKP-------CGSSSAGKIPC 496
                | |   : |     :::   : || | |     || |          |: ||| |
Db        466 --PGGYGPAQQGPSGPGIAASAASAGPGGYGPAQQGPAGYGPGSAVAASAGAGSAGYGPG 523

Qy        497 RSIRDILAQVKPLGPQLADPE 517
                 :|     :|| |:
Db        524 -------SQASAAASRLASPD 537


RESULT 5
US-08-317-844B-4
; Sequence 4, Application US/08317844B
; Patent No. 5989894
;  GENERAL INFORMATION:
;    APPLICANT:  Lewis, Randolph V.
;    APPLICANT:  Xu, Ming
;    APPLICANT:  Hinman, Michael B.
;    TITLE OF INVENTION:  ISOLATED DNA CODING FOR SPIDER SILK
;    TITLE OF INVENTION:  PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
;    TITLE OF INVENTION:  CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
;    NUMBER OF SEQUENCES:  62
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolasch & Birch
;      STREET:  301 No. 5989894th Washington Street
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22046
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/317,844B
;      FILING DATE:  04-OCT-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murphy Jr., Gerald M
;      REGISTRATION NUMBER:  28,977
;      REFERENCE/DOCKET NUMBER:  1447-105P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 241-1300
;      TELEFAX:  (703) 241-2848
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  595 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-317-844B-4

  Query Match             8.9%;  Score 246;  DB 2;  Length 595;
  Best Local Similarity   25.1%;  Pred. No. 1.4e-08;
  Matches  126;  Conservative   56;  Mismatches  217;  Indels  102;  Gaps   23;

Qy         58 GKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQG----------GSAG 104
              | |      |  ||::: |: :|| |   |  |   |      |           | |
Db         98 GYGQGQQGPSGPGSAAAASAAASAESGQQGPGGYGPGQQGPGGYGPGQQGPGGYGPGQQG 157

Qy        105 SFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFS 164
                ||:  :  : :|| |    |  |  | |       | |     |:: :::::|
Db        158 PSGPGSAAAAAAAASGPGQQGPGGYGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAS---- 213

Qy        165 SSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPD 224
                  | |   |       |   | | |         |:|  | : ::
Db        214 ------GPGQQGP-------GGYGPGQQGPGGYGPGQQGLSGPGSAAAA---------- 249

Qy        225 SPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSG--APGVVQGPPCSNGGL-PGK 281
                : ||    | | |    ||          | |   ||  |||    || ||:
Db        250 --AAAGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAAGPGGYGPG-QQGP----GGYGPGQ 302

Qy        282 PCP----PITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVP 337
              |        :   :  | : :||        ||  |  |:  | ||      |  ||
Db        303 QGPSGAGSAAAAAAAGPGQQGLGGYGPGQQGPG-GYGPGQQGPGGY----GP--GSASAA 355

Qy        338 SFAAGP------PISEGKYFSSNPIIPSQSAASSAIAFQPVGTG-GVQLCGGGSTGSKGP 390
              : ||||           |:   | |   | |||::| |  | | | | |  || : | :||
Db        356 AAAAGPGQQGPGGYGPGQQGPSGP--GSASAAAAAAAAGPGGYGPGQQGPGGYAPGQQGP 413

Qy        391 CSPSSSRVPSSSSISSSAGSP--YHP----CGSASQSPCSPPGTGSFSSSSSSQSSGKII 444
                | |   :|:: :::|  |  | |      |  :     | | || :::::: |:|
Db        414 SGPGS----ASAAAAAAAAGPGGYGPGQQGPGGYAPGQQGPSGPGS-AAAAAAASAG--- 465

Qy        445 LQPCGSKSSSSGHPCMSVSSLTLTGGPDG-SPHPDPSAGAKP-------CGSSSAGKIPC 496
                | |   : |     :::   : || | |     || |          |: ||| |
```

```
Db        466 --PGGYGPAQQGPSGPGIAASAASAGPGGYGPAQQGPAGYGPGSAVAASAGAGSAGYGPG 523

Qy        497 RSIRDILAQVKPLGPQLADPE 517
                  :|      :|| |:
Db        524 -------SQASAAASRLASPD 537


RESULT 6
US-09-247-806-2
; Sequence 2, Application US/09247806
; Patent No. 6280747
; GENERAL INFORMATION:
;  APPLICANT: PHILLIPPE, Michel
;  APPLICANT: GARSON, Jean-Claude
;  APPLICANT: ARRAUDEAU, Jean-Pierre
;  TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
;  TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
;  TITLE OF INVENTION: ANALOG
;  FILE REFERENCE: 6388-0365-0
;  CURRENT APPLICATION NUMBER: US/09/247,806
;  CURRENT FILING DATE: 1999-02-11
;  EARLIER APPLICATION NUMBER: FR 98/01614
;  EARLIER FILING DATE: 1998-02-11
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Nephila clavipes
US-09-247-806-2

  Query Match             8.9%;  Score 244.5;  DB 3;  Length 529;
  Best Local Similarity   25.6%;  Pred. No. 1.5e-08;
  Matches  122;  Conservative   53;  Mismatches  209;  Indels   93;  Gaps   22;

Qy         58 GKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQG----------GSAGSF 106
              | |       |  ||::: |: :|| | | |  |       |           | |
Db         98 GYGQGQQGPSGPGSAAAASAAASAESGGPGGYGPGQQGPGGYGPGQQGPGGYGPGQQGPS 157

Qy        107 KPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSS 166
              ||:  :  : :|| |    | .|  | |       | |      |:: :::::|
Db        158 GPGSAAAAAAAAASGPGQQGPGGYGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAS------ 211

Qy        167 SFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSP 226
                | |   |       |   | | |          |:|  | : ::
Db        212 ----GPGQQGP-------GGYGPGQQGPGGYGPGQQGLSGPGSAAAA------------ 247

Qy        227 CSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSG--APGVVQGPPCSNGGL-PGKPC 283
               : ||     | | |     ||           | |   ||  |||    ||  ||:
Db        248 AAAGPGQQGPGGYGPGQQGPSGPGSAAAAAAAAAGPGGYGPG-QQGP----GGYGPGQQG 302

Qy        284 P----PITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSF 339
              |       :   :  | : :||        ||  |  |:  | ||      |  ||    :
Db        303 PSGAGSAAAAAAAGPGQQGLGGYGPGQQGPG-GYGPGQQGPGGY----GP--GSASAAAA 355

Qy        340 AAGP------PISEGKYFSSNPIIPSQSAASSAIAFQPVGTG-GVQLCGGGSTGSKGPCS 392
              ||||          |:   | |   | |||::| |  | | | | |  || : | :||
Db        356 AAGPGQQGPGGYGPGQQGPSGP--GSASAAAAAAAAGPGGYGPGQQGPGGYAPGQQGPSG 413

Qy        393 PSSSRVPSSSSISSSAGSP--YHP----CGSASQSPCSPPGTGSPSSSSSSQSSGKIILQ 446
              | |    :|:: :::|  |  | |    |  :     | | || :::::: |:|
Db        414 PGS----ASAAAAAAAAGPGGYGPGQQGPGGYAPGQQGPSGPGS-AAAAAAASAG----- 463

Qy        447 PCGSKSSSSGHPCMSVSSLTLTGGPDG-SPHPDPSAGAKP-------CGSSSAGKIP 495
              | |    :  |     :::   : || | |     ||  |          |: ||| |
Db        464 PGGYGPAQQGPSGPGIAASAASAGPGGYGPAQQGPAGYGPGSAVAASAGAGSAGYGP 520


RESULT 7
US-09-581-472B-2
; Sequence 2, Application US/09581472B
; Patent No. 6818744
; GENERAL INFORMATION:
;  APPLICANT: FRADE, Raymond
;  TITLE OF INVENTION: p53 REGULATORY PROTEIN CALLED RB18A AND USES THEREOF
;  FILE REFERENCE: P06781US00/BAS
;  CURRENT APPLICATION NUMBER: US/09/581,472B
;  CURRENT FILING DATE:  2000-08-14
;  PRIOR APPLICATION NUMBER: PCT/EP98/08560
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: EP 97403051.2
;  PRIOR FILING DATE: 1997-12-15
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 1566
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-581-472B-2

  Query Match             8.8%;  Score 243.5;  DB 4;  Length 1566;
  Best Local Similarity   27.7%;  Pred. No. 6.2e-08;
  Matches  137;  Conservative   35;  Mismatches  173;  Indels  149;  Gaps   22;

Qy         14 GHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSS 73
              |:|     |:|   ||  :|       |  |     | : |    | |: | |    ||
Db        960 GNGTSNSTLSG---PGLDSKPGKRSRTPSNDG---KSKDKPPKRKKADTEGKS--PSHSS 1011

Qy         74 SGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQ 133
              |    :   |:||  |  |:| |    | |     |                ::|   |:
Db       1012 SNRPFTPPTSTGGSKSPGSAGRSQTPPGVATPPIPKI------------TIQIPKGTVM 1058

Qy        134 LGSSSSHS--GSSG--SHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNP 189
              :|  ||||   |||  | ||| | || ||||   :|:|                  | :
Db       1059 VGKPSSHSQYTSSGSVSSSGSKSHHSHSSSSSSSASTS----------------GKMKS 1101

Qy        190 SQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGG 249
              |:   ||||  :| : ||  |  |:|   ||     | |  ||  |      || | |
Db       1102 SKSEGSSSSKLSSSMYSSQGSSGSSQSKNSSQSGGKPGS-SPITKHG----LSSGSSSTK 1156

Qy        250 QRPVVVVVDQHGSGAPGVVQGPPCSNGGL-PGKPCPPITSVDKSYGGYEVVGGSSDSYLV 308
              :|          | |  :  |  |   : |    ||               | ||
Db       1157 MKP---------QGKPSSLMNPSLSKPNISPSHSRPP--------------GGSDKLAS 1192

Qy        309 PGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIA 368
              |              || |:|  ||  |  |||
Db       1193 P----------------MKPVPGTP--PSSKAKSPISS-------------------- 1212

Qy        369 FQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGT 428
              |:||  :  |          : ||| : ||| : ||       || ||   :|| :
Db       1213 ----GSGGSHMSG----------TSSSSGMKSSSGLGSS--------GSLSQK--TPPSS 1248

Qy        429 GSFSSSSSS-QSSGKIILQPCGSKSSSSG-HPCMSVS-SLT----------LTGGPDGSP 475
              | ::|||  |||  :         || |  |  :  |||           :| || |
Db       1249 NSCTASSSSFSSSGSSMSSSQNQHGSSKGKSPSRNKKPSLTAVIDKLKHGVVTSGPGGED 1308

Qy        476 HPDPSAGAKPCGSS 489
              |  |    ||
Db       1309 PLDGQMGVSTNSSS 1322


RESULT 8
US-09-110-517-2
; Sequence 2, Application US/09110517A
; Patent No. 6248520
; GENERAL INFORMATION:
;  APPLICANT: Roeder, Robert G
;  APPLICANT: Fondell, Joseph D
;  APPLICANT: Yuan, Chao X
;  APPLICANT: Ito, Mitsuhiro
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING NUCLEAR HORMONE
;  TITLE OF INVENTION: RECEPTOR COACTIVATORS AND USES THEREOF
;  FILE REFERENCE: 600-1-224
;  CURRENT APPLICATION NUMBER: US/09/110,517A
;  CURRENT FILING DATE: 1998-07-06
;  NUMBER OF SEQ ID NOS: 51
```

```
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1581
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-110-517-2

  Query Match             8.8%;  Score 243.5;  DB 3;  Length 1581;
  Best Local Similarity   27.7%;  Pred. No. 6.3e-08;
  Matches  137;  Conservative   35;  Mismatches  173;  Indels  149;  Gaps   22;

Qy         14 GHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSS 73
              |:|     |:|   ||  :|       | |     | : |    | |: | |     ||
Db        975 GNGTSNSTLSG---PGLDSKPGKRSRTPSNDG---KSKDKPPKRKKADTEGKS--PSHSS 1026

Qy         74 SGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQ 133
              |    :  |:|| |  |:| |     | |     |              ::|   |:
Db       1027 SNRPFTPPTSTGGSKSPGSAGRSQTPPGVATPPIPKI------------TIQIPKGTVM 1073

Qy        134 LGSSSSHS--GSSG--SHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNP 189
              :|  ||||   |||  | ||| | || ||||   :|:|              | :
Db       1074 VGKPSSHSQYTSSGSVSSSGSKSHHSHSSSSSSSASTS----------------GKMKS 1116

Qy        190 SQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGG 249
              |:   |||| :| : ||  |  |:|   ||      | |  ||  |        || | |
Db       1117 SKSEGSSSSKLSSSMYSSQGSSGSSQSKNSSQSGGKPGS-SPITKHG----LSSGSSSTK 1171

Qy        250 QRPVVVVVDQHGSGAPGVVQGPPCSNGGL-PGKPCPPITSVDKSYGGYEVVGGSSDSYLV 308
               :|          | |  |  |   : |    ||              | ||
Db       1172 MKP---------QGKPSSLMNPSLSKPNISPSHSRPP----------GGSDKLAS 1207

Qy        309 PGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIA 368
              |                 || |:|  ||  |  |||
Db       1208 P-----------------MKPVPGTP--PSSKAKSPISS-------------------- 1227

Qy        369 FQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGT 428
                 |:||  : |          : ||| : ||| : ||        || ||  :|| :
Db       1228 ----GSGGSHMSG----------TSSSSGMKSSSGLGSS--------GSLSQK--TPPSS 1263

Qy        429 GSFSSSSSS-QSSGKIILQPCGSKSSSSG-HPCMSVS-SLT----------LTGGPDGSP 475
               | ::|||| |||  :         || | |  :  |||          :| || |
Db       1264 NSCTASSSSFSSSGSSMSSSQNQHGSSKGKSPSRNKKPSLTAVIDKLKHGVVTSGPGGED 1323

Qy        476 HPDPSAGAKPCGSS 489
              |   |   ||
Db       1324 PLDGQMGVSTNSSS 1337


RESULT 9
US-09-949-016-10366
; Sequence 10366, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10366
;   LENGTH: 1259
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-10366

  Query Match             8.8%;  Score 243;  DB 4;  Length 1259;
  Best Local Similarity   28.1%;  Pred. No. 5.2e-08;
  Matches  132;  Conservative   26;  Mismatches  247;  Indels   64;  Gaps    8;

Qy         39 SDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIA 98
              ||   :     | |    :  |||  |  | || | ||  |: ||    |  || || :
Db        655 SDSSDNSDSSDSSNSSNSSDSSDSSDSSDSSSSSDSSSSSDSSNSSDSSDSSDSSNSSES 714

Qy         99 QGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSS 158
                 |    :  | | |: | |    :| ||    || | || |   | ||:|| |
Db        715 SDSSDSSDSDSSDSSDSSNSNSSDSDSSNSSDSSDSSDSSDSSNSSDSSDSSDSSNSSDS 774

Qy        159 SSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVS----------S 206
              |   || |    | |  :||:|   | |    ||:|| :|  |            |
Db        775 SDSSDSSDSSDSSNSS--DSNDSS-----NSSDSSDSSNSSDSSNSSDSSDSSDSDS 827

Qy        207 SGQSVSSNQRPC----SSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGS 262
              |  | |||         |||  ||  |    | |     ||:|           |
Db        828 SNSSDSSNSSDSSDSNSSDSSDSSDSSDSSDSSDSSNRSDSSNSSDSSD----------SS 877

Qy        263 GAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVG 322
              :          |:     :     | | |          ||||          |
Db        878 DSSNSSDSSDSSDSSDSNESSNSSDSSDSS--------NSSDSDSSDSSNSSDSSDSSNS 929

Qy        323 YFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGG 382
                : |:          | ::     |      |||     | |  ||
Db        930 SDSSESSNSSDSNSSSDSSNSSDSSDSSDSSNSSDSSNSGDSS----------------N 973

Qy        383 GSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSQSSGK 442
              |  |    | ||    || |  ||  |        :| |  |   : |  ||:|| ||
Db        974 SSDSSDSNSSDSSDSSNSSDSSDSSDSSDSSDSSDSSNSSDSSDSSDSSDSSNSSDSSN- 1032

Qy        443 IILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSA 491
                  | || |     |  |   :   | |  |  |   :   ||| :
Db       1033 ------SSDSSDSSDSSDSSDSSDSSNSSDSSDSSDSSDSSDSSGSSDS 1075


RESULT 10
US-09-538-092-997
; Sequence 997, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
;  APPLICANT: Giot, Loic
;  APPLICANT:  Mansfield, Traci A.
;  TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
;  FILE REFERENCE: 15966-542
;  CURRENT APPLICATION NUMBER: US/09/538,092
;  CURRENT FILING DATE:  2000-03-29
;  PRIOR APPLICATION NUMBER: 60/127,352
;  PRIOR FILING DATE: 1999-04-01
;  PRIOR APPLICATION NUMBER: 60/178,965
;  PRIOR FILING DATE: 2000-02-01
;  NUMBER OF SEQ ID NOS: 1387
;  SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 997
;   LENGTH: 316
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Polypeptide Accession Number P23490
US-09-538-092-997

  Query Match             8.8%;  Score 241.5;  DB 4;  Length 316;
  Best Local Similarity   25.1%;  Pred. No. 1.3e-08;
  Matches  115;  Conservative   19;  Mismatches  150;  Indels  175;  Gaps   21;

Qy         43 KDPTRITSPNDP--CLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQG 100
              | ||     | |  |:   |    |     |:   |          |||||||| | |   |
Db          6 KQPT----PQPPVDCVKTSGGGGG---GGGTGGGGCGFFGGGGSGGGGSSGSGCGYSGGGG 58

Qy        101 GSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSS 160
```

```
Db         59 YSGGGCGGG------SSGGGGGGGIGGCGGGS--GGSVKYSGGGGSGGGSGCFSSGGG- 109

Qy        161 FQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSS 220

Db        110 ---GSGCFSSGGG----------------------GSSGGGSGCFSSG------------ 132

Qy        221 DIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPC---SNGG 277

Db        133 -------GGGS---------------SGG----------GSG---------CFSSGGGG 150

Qy        278 LPGK--PCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPG 335

Db        151 FSGQAVQC-------QSYGGVS-SGGSSGG----------------------------G 173

Qy        336 VPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSS 395

Db        174 SGCFSSG-----------------GGGGSVCGYSGGGSGGGSGCGGGSSGGSG------ 209

Qy        396 SRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSS 455

Db        210 SGYVSSQQVTQTSCAPQPSYGGGSSGGGGSGGSGCFSSGGGGGSSG------CGGGSSGI 263

Qy        456 GHPC-MSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGK 493

Db        264 GSGCIISGGGSVCGGGSSGG-----GGGGSSVGGSGSGK 297


RESULT 11
US-09-248-796A-23039
; Sequence 23039, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 23039
;   LENGTH: 461
;   TYPE: PRT
;   ORGANISM: Candida albicans
US-09-248-796A-23039

  Query Match             8.7%;  Score 239.5;  DB 4;  Length 461;
  Best Local Similarity   27.6%;  Pred. No. 2.7e-08;
  Matches  121;  Conservative   60;  Mismatches  214;  Indels   43;  Gaps   12;

Qy         45 PTRITSPNDPCLTGKGDSSG--FSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGS 102

Db         15 PVKSALPSEVSSSSEVSSSSQVISSSEVVSSSSEVVSSSSEVSSSSEVSSSSEVVSSSSE 74

Qy        103 AGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQ 162

Db         75 VSSSSEVSSSSEVSSSSQVTSSSEIVSSSSEVSSSSSEVVSSSSEVSSSSEVVSSSSEVS 134

Qy        163 FS---SSSFQVGNGSALPTND--NSYRGILNPSQPGQSSS--SSQTSGVSSSGQSVSSNQ 215

Db        135 SSSEVSSSSEVSSSSEVSSSSEVSSSSQVISSSEVVSSSSEVSSSSEVVSSSSEVSSSSE 194

Qy        216 RPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSN 275

Db        195 VSSSSEV-----SSSSQVTSSSEIVSSSSEVSSSSSEVVSSSSEVSSSSEVVSSSSEVSS 249

Qy        276 GGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPG 335

Db        250 SS-EVSSSSEVSSSSQVTSSSEIVSSSSE------VSSSSSEV-----------VSSSSE 291
```

```
Qy        336 VPSFAAGPPISEGKYFSSNPIIPSQSAASSA---IAFQPVGTGGVQLCGGGSTGSKGPCS 392

Db        292 VSS--SSEVVSSSSEVSSSSEVSSSSEVSSSSQVISSSEVVSSSSEVVSSSSEVSSSSEV 349

Qy        393 PSSSRVPSSSSISSSA--GSPYHPCGS----ASQSPCSPPGTGSFSSSSSSQSSGKIILQ 446

Db        350 SSSSEVSSSSEVSSSSEVSSSSQVTSSSEIVSSSSEVSSSSSEVVSSSSEVSSSSEVVSS 409

Qy        447 PCGSKSSSSGHPCMSVSS 464

Db        410 SSEVSSSSEVSSSSEVSS 427


RESULT 12
US-09-411-067C-5
; Sequence 5, Application US/09411067C
; Patent No. 6576610
; GENERAL INFORMATION:
;  APPLICANT: NUVAS, LLC
;  APPLICANT:  HOUSTON, L.L.
;  TITLE OF INVENTION: USE OF A CONTEXT-DEPENDENT FUNCTIONAL ENTITY TO ENHANCE THE EFFIC
;  TITLE OF INVENTION:  AN AGENT
;  FILE REFERENCE: NUVAS1140
;  CURRENT APPLICATION NUMBER: US/09/411,067C
;  CURRENT FILING DATE:  2002-05-07
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 100
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Spacer element
;   FEATURE:
;   NAME/KEY: REPEAT
;   LOCATION: (1)..(5)
;   OTHER INFORMATION: Spacer peptide variation; sequence repeated "n" times, where n =
;   OTHER INFORMATION: 1-20
US-09-411-067C-5

  Query Match             8.2%;  Score 225.5;  DB 4;  Length 100;
  Best Local Similarity   46.9%;  Pred. No. 3.9e-08;
  Matches   67;  Conservative    4;  Mismatches   29;  Indels   43;  Gaps    3;

Qy         66 SSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSL 125

Db          1 SSSSGSSSSGSSSSGSSSSGSSSSGSSSSGSSSSGSS--------------SSGSSSSG 45

Qy        126 QGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRG 185

Db         46 SSSSGSSSSGSSSSGSSSSGSSSSGSSSSGSSSSG---SSSS----------------- 84

Qy        186 ILNPSQPGQSSSSSQTSGVSSSG 208

Db         85 -------GSSSSGSSSSGSSSSG 100


RESULT 13
US-09-538-092-330
; Sequence 330, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
;  APPLICANT: Giot, Loic
;  APPLICANT:  Mansfield, Traci A.
;  TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
;  FILE REFERENCE: 15966-542
;  CURRENT APPLICATION NUMBER: US/09/538,092
;  CURRENT FILING DATE:  2000-03-29
;  PRIOR APPLICATION NUMBER: 60/127,352
;  PRIOR FILING DATE: 1999-04-01
;  PRIOR APPLICATION NUMBER: 60/178,965
;  PRIOR FILING DATE: 2000-02-01
```

```
;  NUMBER OF SEQ ID NOS: 1387
;  SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 330
;   LENGTH: 1306
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Polypeptide Accession Number YGR014W
US-09-538-092-330

  Query Match             8.1%;  Score 223.5;  DB 4;  Length 1306;
  Best Local Similarity   25.7%;  Pred. No. 1.1e-06;
  Matches  134;  Conservative   63;  Mismatches  219;  Indels  105;  Gaps   20;

Qy         18 MALLLAGLLLPGT------LAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGS 71
              |: ||:   |  |      || |  : ||   |    :: : | |:    :|  ||:|
Db        427 MSSLLSTTSLQSTTLDSSSLASSSASSSD-LTDYGVSSTASIPLLSASEQASTSSSFSVV 485

Qy         72 SSSGSSISSARSSGGGS-------SGSSSGSSIAQGGSAGS--------FKPGTGYSQVS 116
              | | | : |  ||   |       | | | :|:  |||: :       :   || |: |
Db        486 SPSVSFVPSQSSSDVASTSAPSVVSSSFSYTSLQAGGSSMTNPSSSTIVYSSSTGSSEES 545

Qy        117 YSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSS----SSSFQFSSSSFQVGN 172
              :| : ::| |:| :   |:  |   |: |    | : |:|    |||   :|     |
Db        546 AASTASATLSGSSSTYMAGNLQSQPPSTSSLLSESQATSTSAVLASSSVSTTSPYTTAGG 605

Qy        173 GS----------ALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSN----QRPC 218
              ||          :  |:  ||       |   :|||  | |  :| |  |::    | |
Db        606 ASTEASSLISSTSAETSQVSYSQSTTALQTSSFASSSTTEGSETSSQGFSTSSVLVQMPS 665

Qy        219 SSDIPDSP----------CSGGPIVSHSG----------PYIPSSHSVSG-GQRPVVVV 256
              |    ||           |    :| :|          |  || |||    ||
Db        666 SISSEFSPSQTTTQMNSASSSSQYTISSTGILSQVSDTSVSYTTSSSSVSQVSDTPVSYT 725

Qy        257 VDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKG 316
                | :   |  |                  |  |       |  :  ||    : |:
Db        726 T---SSSSVSQVSDTP-------------VSYTTSSSSVSQVSDTPVSYTTSSSSVSQV 768

Qy        317 KIYPVGYFTKENPVK--GSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASS----AIAFQ 370
              || | |  : |       ||| ::   :|:    |  | :|| |: ||    : : |
Db        769 SDTPVSYTTSSSSVSQVSDTSVPSTSSRSSVSQ---VSDTP-VPSTSSRSSVSQTSSSLQ 824

Qy        371 PVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSS---AGSPYHPCGSASQSPCSPPG 427
              |  |   :     :  : |  | |||    :| |:||     | ||   |   :  :
Db        825 PTTTSSQRF----TISTHGALSESSSVSQQASEITSSINATASEYH---SIQTTAATQST 877

Qy        428 TGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLT 468
              | ||: ::|| :|  :        ::|:  |    ||| ||
Db        878 TLSFTDANSSSASAPL------EVATSTPTPSSKASSLLLT 912


RESULT 14
US-09-538-092-1005
; Sequence 1005, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
;  APPLICANT: Giot, Loic
;  APPLICANT:  Mansfield, Traci A.
;  TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
;  FILE REFERENCE: 15966-542
;  CURRENT APPLICATION NUMBER: US/09/538,092
;  CURRENT FILING DATE:  2000-03-29
;  PRIOR APPLICATION NUMBER: 60/127,352
;  PRIOR FILING DATE: 1999-04-01
;  PRIOR APPLICATION NUMBER: 60/178,965
;  PRIOR FILING DATE: 2000-02-01
;  NUMBER OF SEQ ID NOS: 1387
;  SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 1005
;   LENGTH: 1970
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Polypeptide Accession Number P24928
US-09-538-092-1005

  Query Match             8.0%;  Score 221;  DB 4;  Length 1970;
  Best Local Similarity   27.1%;  Pred. No. 2.5e-06;
  Matches  144;  Conservative   48;  Mismatches  215;  Indels  124;  Gaps   26;

Qy          7 PWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTR--------ITSPNDPCLTG 58
              | :|  |  ||           |:    :|  | |   |             ||:     |
Db       1491 PGLGAAGPTGMFF---------GSAPSPMGGIS-PAMTPWNQGATPAYGAWSPS----VG 1536

Qy         59 KG---DSSGFSSYSGSSSSGSS--ISSARSSGGGSSGSSSGSSIAQGGSAGSPKPGTGYS 113
               |    ::|||  : | :|| |   | | |   || ||   ||        |: |    :
Db       1537 SGMTPGAAGFSPSAASDASGFSPGYSPAWSPTPGSPGSPGPSSPYIPSPGGAMSPSYSPT 1596

Qy        114 QVSYS--SGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVG 171
                :|   | | : |  | |    :| |:| :| |:| :| ::| :| |:  :| |:
Db       1597 SPAYEPRSPGGYTPQSPSYSP---TSPSYSPTSPSYSPTSPNYSPTSPSYSPTSPSYSPT 1653

Qy        172 NGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCS-SDIPDSPCSG- 229
              : |  ||:  ||     :|: |  | :|   |  | |     | :  | | |  | ||
Db       1654 SPSYSPTSP-SY----SPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSP 1708

Qy        230 -GPIVSHSGP-YIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKP----- 282
               |  | : | | |:| | |                |   | |   |  |
Db       1709 TSPSYSPTSPSYSPTSPSYS----------------PTSPSYSPTSPSYSPTSPNYSPT 1751

Qy        283 CPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAG 342
              |  |    ||         :| ||      |           |  :|   ||  ||::
Db       1752 SPNYTPTSPSY------SPTSPSYSPTSPNY----------TPTSP-NYSPTSPSYSPT 1793

Qy        343 PPISEGKYFSSNPIIPSQSAASSAIAFQ-PVGTGGVQLCGGGSTGSKGPCSPSSS----- 396
              |       | :|  || | :|      | | |          |: |  | ||| |
Db       1794 SP-------SYSPTSPSYSPSSPRYTPQSPTYT--------PSSPSYSPSSPSYSPTSPK 1838

Qy        397 RVPSSSSISSSA-----------GSPYHPCGSASQSPCSP---PGTGSFSSSSSSQSSG 441
              |:| | | |:             || :   |   || ||   | |  :| :| : |
Db       1839 YTPTSPSYSPSSPEYTPTSPKYSPTSPKYSPTSPKYSPTSPTYSPTTPKYSPTSPTYSPT 1898

Qy        442 KIILQPCGSKSS------SSGHPCMSVSSLTLT-GGPDGSPHPDPSAGAKP 485
              :  |   | |       |   | | :| | :   | || :   | |  |
Db       1899 SPVYTPTSPKYSPTSPTYSPTSPKYSPTSPTYSPTSPKGSTYSPTSPGYSP 1949


RESULT 15
US-09-538-092-1160
; Sequence 1160, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
;  APPLICANT: Giot, Loic
;  APPLICANT:  Mansfield, Traci A.
;  TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
;  FILE REFERENCE: 15966-542
;  CURRENT APPLICATION NUMBER: US/09/538,092
;  CURRENT FILING DATE:  2000-03-29
;  PRIOR APPLICATION NUMBER: 60/127,352
;  PRIOR FILING DATE: 1999-04-01
;  PRIOR APPLICATION NUMBER: 60/178,965
;  PRIOR FILING DATE: 2000-02-01
;  NUMBER OF SEQ ID NOS: 1387
;  SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 1160
;   LENGTH: 1475
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Polypeptide Accession Number P49790
```

US-09-538-092-1160

Query Match 8.0%; Score 219.5; DB 4; Length 1475;
Best Local Similarity 20.9%; Pred. No. 2.3e-06;
Matches 146; Conservative 68; Mismatches 232; Indels 253; Gaps 27;

```
Qy         28 PGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSS---------------YSGSS 72
              |||  |   | :  :    :|: :  |     | : ||                  | ::
Db        750 PGTCVKRALTLTVVSESAETMTASSSSCTVTTG-TLGFGDKFKRPIGSWECSVCCVSNNA 808

Qy         73 SSGSSISSARSSGGGSSGSSSGS----SIAQGGSAG------------------------ 104
                :|      | |  :|| |    |:  ||| |
Db        809 EDNKCVSCMSEKPGSSVPASSSSTVPVSLPSGGSLGLEKFKKPEGSWDCELCLVQNKADS 868

Qy        105 -------SFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSS-----------GS 146
                  | ||||      : : | ||   || | : | ||| || |            |
Db        869 TKCLACESAKPGTKSGFKGFDTSSSSSNSAASSSFKFGVSSSSSGPSQTLTSTGNFKFGD 928

Qy        147 HSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTS---- 202
               |      || | |     |  |:       |  |  : |: : |:| :    |:
Db        929 QGGFKIGVSSDSGSINPMSEGFKFSK----PIGDFKF-GVSSESKPEEVKKDSKNDNFKF 983

Qy        203 -----------------GVSSSGQSVSSNQRPCSSD---------IPDSPCSGGPIVSHS 236
                          |||: ||      : | ||          |  :|      ||:
Db        984 GLSSGLSNPVSLTPFQFGVSNLGQEEKKEELPKSSSAGFSFGTGVINSTPAPANTIVTSE 1043

Qy        237 -----------------GPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVV----------- 268
                              |:   :       : |         |:  |  :
Db       1044 NKSSFNLGTIETKSASVAPFTCKTSEAKKEEMPATKGGFSFGNVEPASLPSASVFVLGRT 1103

Qy        269 ----QGPPCSNGGLPGK-------PCPPITSV-----------DKSYGGYEVVGGSSDSY 306
                 | |  |   : ||         | |: |             ||   : :   |
Db       1104 EEKQQEPVTSTSLVFGKKADNEEPKCQPVFSFGNSEQTKDENSSKSTFSFSMTKPSEKES 1163

Qy        307 LVP-----------GMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAG------------P 343
               |           |  :|   ||  |:  |     |   |: :||             |
Db       1164 EQPAKATFAFGAQTSTTADQGAAKPV--FSFLNNSSSSSSTPATSAGGGIFGSSTSSSNP 1221

Qy        344 PISEGKY-FSSNPIIPS------QSAASSAIAFQPVGTGGVQLCGGGSTGSK------GP 390
              |::  :  ||||:  |      :|: | :: |    :   :|    |||:        ||
Db       1222 PVATFVFGQSSNPVSSSAFGNTAESSTSQSLLF----SQDSKLATTSSTGTAVTPFVFGP 1277

Qy        391 CSPSSSRVPS-----SSSISSSAGSPY-----------HPCGSASQSPC----------S 424
              : |::   |     ::: |||||| :            |   |:|:|             :
Db       1278 GASSNNTTTSGFGFGATTTSSSAGSSFVFGTGPSAPSASPAFGANQTPTFGQSQGASQPN 1337

Qy        425 PPGTGSFSSSSSSQSSGKIILQPC----GSKSSSSGHPCM-------------------- 460
              |||  || |||::   :|    ||      |: ||||  |
Db       1338 PPGFGSISSSTALFPTGS---QPAPPTFGTVSSSSQPPVFGQQPSQSAFGSGTTPNSSSA 1394

Qy        461 -----SVSSLTLTGGPD------GSPHPDPSAGAKPCGS 488
                  | ::   |          |:    |:| |:| ||
Db       1395 FQFGSSTTNFNFTNNSPSGVFTFGANSSTPAASAQPSGS 1433
```

Search completed: August 2, 2005, 09:13:58
Job time : 47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 2, 2005, 09:12:29 ; Search time 163 Seconds
(without alignments)
1264.433 Million cell updates/sec

Title: US-10-677-865-1
Perfect score: 2752
Sequence: 1 MGSSRAPWMGRVGGHGMMAL.....GPQLADPEVFLPQGELLDSP 529

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1745140 seqs, 389608008 residues

Total number of hits satisfying chosen parameters: 1745140

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2752	100.0	529	15	US-10-316-175-15	Sequence 15, Appl
2	2752	100.0	529	17	US-10-677-865-1	Sequence 1, Appli
3	2734	99.3	529	17	US-10-883-805-6	Sequence 6, Appli
4	2707	98.4	529	14	US-10-204-884-3	Sequence 3, Appli
5	2697	98.0	529	14	US-10-164-230-61	Sequence 61, Appl
6	2505	91.0	486	15	US-10-316-175-12	Sequence 12, Appl
7	2505	91.0	486	16	US-10-723-860-4324	Sequence 4324, Ap
8	2505	91.0	486	17	US-10-883-805-16	Sequence 16, Appl
9	1106	40.2	219	14	US-10-029-386-34046	Sequence 34046, A
10	717	26.1	146	15	US-10-316-175-71	Sequence 71, Appl
11	533	19.4	111	15	US-10-316-175-72	Sequence 72, Appl
12	421	15.3	81	15	US-10-316-175-73	Sequence 73, Appl
13	310	11.3	58	15	US-10-316-175-70	Sequence 70, Appl
14	298	10.8	55	15	US-10-316-175-69	Sequence 69, Appl
15	292.5	10.6	528	10	US-09-840-746-20	Sequence 20, Appl
16	276	10.0	668	16	US-10-437-963-134726	Sequence 134726,
17	266	9.7	258	15	US-10-104-047-3034	Sequence 3034, Ap
18	260	9.4	384	16	US-10-479-638-19	Sequence 19, Appl
19	257	9.3	1459	15	US-10-369-493-6418	Sequence 6418, Ap
20	253.5	9.2	1217	14	US-10-311-406-2	Sequence 2, Appli
21	248.5	9.0	3507	15	US-10-369-493-5784	Sequence 5784, Ap
22	248	9.0	627	16	US-10-406-832-28	Sequence 28, Appl
23	248	9.0	627	17	US-10-893-588-28	Sequence 28, Appl
24	248	9.0	1367	9	US-09-801-368-108	Sequence 108, App
25	244.5	8.9	529	9	US-09-861-597-2	Sequence 2, Appli
26	244.5	8.9	529	17	US-10-887-100-2	Sequence 2, Appli
27	244.5	8.9	1253	14	US-10-363-798-2	Sequence 2, Appli
28	244	8.9	624	16	US-10-479-638-12	Sequence 12, Appl
29	243.5	8.8	584	15	US-10-282-122A-50604	Sequence 50604, A
30	243.5	8.8	1566	14	US-10-177-293-366	Sequence 366, App
31	243.5	8.8	1566	15	US-10-435-696-30	Sequence 30, Appl
32	243.5	8.8	1566	15	US-10-425-970-2	Sequence 2, Appli
33	242	8.8	465	16	US-10-437-963-174113	Sequence 174113,
34	241.5	8.8	316	16	US-10-723-860-2777	Sequence 2777, Ap
35	239.5	8.7	518	15	US-10-282-122A-50634	Sequence 50634, A
36	239	8.7	1814	17	US-10-488-056-53	Sequence 53, Appl
37	238.5	8.7	406	16	US-10-451-467A-262	Sequence 262, App
38	238	8.6	399	17	US-10-488-056-41	Sequence 41, Appl
39	236.5	8.6	551	15	US-10-425-114-43482	Sequence 43482, A
40	235.5	8.6	420	16	US-10-437-963-201378	Sequence 201378,
41	234	8.5	465	16	US-10-425-115-203464	Sequence 203464,
42	233	8.5	343	14	US-10-148-306-21	Sequence 21, Appl
43	231	8.4	903	14	US-10-156-761-11093	Sequence 11093, A
44	230	8.4	497	16	US-10-437-963-125004	Sequence 125004,
45	229.5	8.3	566	16	US-10-425-115-336849	Sequence 336849,

ALIGNMENTS

RESULT 1
US-10-316-175-15
; Sequence 15, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
; APPLICANT: WELLA AG
; APPLICANT: Pras, Elon
; APPLICANT: Levy-Nissenbaum, Etgar
; APPLICANT: Frydman, Moshe
; APPLICANT: Thederahn, Ted
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING HYPOTRICHOSIS SIMPLEX
; FILE REFERENCE: 532862000100
; CURRENT APPLICATION NUMBER: US/10/316,175
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/338,188
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-316-175-15

Query Match 100.0%; Score 2752; DB 15; Length 529;
Best Local Similarity 100.0%; Pred. No. 4.4e-150;
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60

Qy         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy         121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180

Qy         181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240

Qy         241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300

Qy         301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360

Qy         361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420

Qy         421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480

Qy         481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 2
US-10-677-865-1
; Sequence 1, Application US/10677865
; Publication No. US20050112150A1
; GENERAL INFORMATION:
;  APPLICANT: SERRE, Guy Bruno Rene
;  APPLICANT:  SIMON, Michel
;  APPLICANT:  WEBER-VIVAT, Marina
;  TITLE OF INVENTION: POLYPEPTIDE EXPRESSED IN THE HORNY LAYER OF EPIDERMIS
;  TITLE OF INVENTION:  AND USE THEREOF
;  FILE REFERENCE: 016800-336
;  CURRENT APPLICATION NUMBER: US/10/677,865
;  CURRENT FILING DATE:  2003-10-03
;  PRIOR APPLICATION NUMBER: US/09/381,656
;  PRIOR FILING DATE: 1999-12-29
;  PRIOR APPLICATION NUMBER: FR 97/03899
;  PRIOR FILING DATE: 1997-03-28
;  PRIOR APPLICATION NUMBER: FR 97/11317
;  PRIOR FILING DATE: 1997-09-11
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Human
US-10-677-865-1

  Query Match            100.0%;  Score 2752;  DB 17;  Length 529;
  Best Local Similarity  100.0%;  Pred. No. 4.4e-150;
  Matches  529;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60

Qy          61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy         121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180

Qy         181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240

Qy         241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300

Qy         301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360

Qy         361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420

Qy         421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480

Qy         481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 3
US-10-883-805-6
; Sequence 6, Application US/10883805
; Publication No. US20050048612A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE PHARMACEUTICALS, INC.
;  APPLICANT:  TANG, Y. Tom
;  APPLICANT:  LAL, Preeti
;  APPLICANT:  CORLEY, Neil C.
;  APPLICANT:  GUEGLER, Karl J.
;  APPLICANT:  PATTERSON, Chandra
;  APPLICANT:  BAUGHN, Mariah R.
;  APPLICANT:  YUE, Henry
;  TITLE OF INVENTION: HUMAN EPIDERMAL PROTEINS
;  FILE REFERENCE: PF-0567 PCT
;  CURRENT APPLICATION NUMBER: US/10/883,805
;  CURRENT FILING DATE:  2004-07-06
;  PRIOR APPLICATION NUMBER: 09/123,531
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: 09/206,817
;  PRIOR FILING DATE: 1998-12-07
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-12-07
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PERL Program
; SEQ ID NO 6
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte Clone No: 3325323
US-10-883-805-6

  Query Match            99.3%;  Score 2734;  DB 17;  Length 529;
  Best Local Similarity  99.2%;  Pred. No. 4.7e-149;
  Matches  525;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy           1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
                ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MGSSRAPWMGRVGGHGMLALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60

Qy          61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180

Qy        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        181 NSYRGILNPSQPGQSSSSSQTFGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240

Qy        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300

Qy        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360

Qy        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||:| ||||||||||
Db        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGLPYHPCGSASQ 420

Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480

Qy        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 4
US-10-204-884-3
; Sequence 3, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (18)..(18)
;   OTHER INFORMATION: The 'Xaa' at location 18 stands for Met, or Leu.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (56)..(56)
;   OTHER INFORMATION: The 'Xaa' at location 56 stands for Leu, or Phe.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (143)..(143)
;   OTHER INFORMATION: The 'Xaa' at location 143 stands for Ser, or Asn.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (202)..(202)
;   OTHER INFORMATION: The 'Xaa' at location 202 stands for Ser, or Phe.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (401)..(401)
;   OTHER INFORMATION: The 'Xaa' at location 401 stands for Gly, or Ser.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (408)..(408)
;   OTHER INFORMATION: The 'Xaa' at location 408 stands for Ala, or Ser.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (409)..(409)
;   OTHER INFORMATION: The 'Xaa' at location 409 stands for Gly, or Val.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (410)..(410)
;   OTHER INFORMATION: The 'Xaa' at location 410 stands for Ser, or Leu.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (527)..(527)
;   OTHER INFORMATION: The 'Xaa' at location 527 stands for Asp, or Asn.
US-10-204-884-3

  Query Match             98.4%;  Score 2707;  DB 14;  Length 529;
  Best Local Similarity   98.3%;  Pred. No. 1.7e-147;
  Matches  520;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              |||||||||||||||| ||||||||||||||||||||||||||||||||||| ||||
Db          1 MGSSRAPWMGRVGGHGMXALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCXTGKG 60

Qy         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 SGSSLQGASGSSQLGSSSSHSGXSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180

Qy        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        181 NSYRGILNPSQPGQSSSSSQTXGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240

Qy        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300

Qy        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360

Qy        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
              |||||||||||||||||||||||||||||||||||||||||||| ||||||   |||||||||
Db        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSXSSISSSXXXPYHPCGSASQ 420

Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480

Qy        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
              |||||||||||||||||||||||||||||||||||||||||||||| ||
Db        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLXSP 529


RESULT 5
US-10-164-230-61
; Sequence 61, Application US/10164230
; Publication No. US20030170652A1
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
```

```
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 143
;   OTHER INFORMATION: Xaa = Ser or Asn
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 151
;   OTHER INFORMATION: Xaa = Ser or Cys
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 152
;   OTHER INFORMATION: Xaa = Ser or Arg
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 202
;   OTHER INFORMATION: Xaa = Ser or Phe
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 282
;   OTHER INFORMATION: Xaa = Pro or Thr
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 401
;   OTHER INFORMATION: Xaa = Ser or Gly
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 408
;   OTHER INFORMATION: Xaa = Ala or Ser
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 409
;   OTHER INFORMATION: Xaa = Gly or Val
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 410
;   OTHER INFORMATION: Xaa = Ser or Leu
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 453
;   OTHER INFORMATION: Xaa = Ser or Asn
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 527
;   OTHER INFORMATION: Xaa = Asp or Asn
US-10-164-230-61

  Query Match             98.0%;  Score 2697;  DB 14;  Length 529;
  Best Local Similarity   97.9%;  Pred. No. 6.3e-147;
  Matches  518;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60

Qy         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
              |||||||||||||||||||| |||||||  ||||||||||||||||||||||||||||||
Db        121 SGSSLQGASGSSQLGSSSSHSGXSGSHSGSXXSHSSSSSSFQFSSSSFQVGNGSALPTND 180

Qy        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        181 NSYRGILNPSQPGQSSSSSQTXGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240

Qy        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKXCPPITSVDKSYGGYEVVG 300

Qy        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360

Qy        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
              |||||||||||||||||||||||||||||||||||||| ||||||   ||||||||||
Db        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSXSSISSSXXXPYHPCGSASQ 420

Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSXSSGHPCMSVSSLTLTGGPDGSPHPDPS 480

Qy        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
              |||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLXSP 529


RESULT 6
US-10-316-175-12
; Sequence 12, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
;   LENGTH: 486
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-12

  Query Match             91.0%;  Score 2505;  DB 15;  Length 486;
  Best Local Similarity   99.6%;  Pred. No. 6e-136;
  Matches  482;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         17 MMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGS 76
              |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGS 60

Qy         77 SISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGS 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGS 120

Qy        137 SSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSS 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSS 180

Qy        197 SSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVV 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVV 240

Qy        257 VDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKG 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKG 300

Qy        317 KIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGG 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGG 360
```

```
Qy        377 VQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSS 436
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        361 VQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGSPYHPCGSASQSPCSPPGTGSFSSSSS 420

Qy        437 SQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPC 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPC 480

Qy        497 RSIR 500
              ||||
Db        481 RSIR 484


RESULT 7
US-10-723-860-4324
; Sequence 4324, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Ginsburg, Wendy M.
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
;  TITLE OF INVENTION:  Methods for Screening for Soft Tissue Sarcoma Modulators
;  FILE REFERENCE: 05882.0193.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/723,860
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,739
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 8393
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4324
;   LENGTH: 486
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-723-860-4324

  Query Match             91.0%;  Score 2505;  DB 16;  Length 486;
  Best Local Similarity   99.6%;  Pred. No. 6e-136;
  Matches  482;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         17 MMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGPSSYSGSSSSGS 76
              |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGPSSYSGSSSSGS 60

Qy         77 SISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGS 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGS 120

Qy        137 SSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSS 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSS 180

Qy        197 SSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVV 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVV 240

Qy        257 VDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKG 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKG 300

Qy        317 KIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGG 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGG 360

Qy        377 VQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSS 436
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        361 VQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGSPYHPCGSASQSPCSPPGTGSFSSSSS 420

Qy        437 SQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPC 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPC 480

Qy        497 RSIR 500
              ||||
Db        481 RSIR 484


RESULT 8
US-10-883-805-16
; Sequence 16, Application US/10883805
; Publication No. US20050048612A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE PHARMACEUTICALS, INC.
;  APPLICANT:  TANG, Y. Tom
;  APPLICANT:  LAL, Preeti
;  APPLICANT:  CORLEY, Neil C.
;  APPLICANT:  GUEGLER, Karl J.
;  APPLICANT:  PATTERSON, Chandra
;  APPLICANT:  BAUGHN, Mariah R.
;  APPLICANT:  YUE, Henry
;  TITLE OF INVENTION: HUMAN EPIDERMAL PROTEINS
;  FILE REFERENCE: PF-0567 PCT
;  CURRENT APPLICATION NUMBER: US/10/883,805
;  CURRENT FILING DATE:  2004-07-06
;  PRIOR APPLICATION NUMBER: 09/123,531
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: 09/206,817
;  PRIOR FILING DATE: 1998-12-07
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-12-07
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PERL Program
; SEQ ID NO 16
;   LENGTH: 486
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: GenBank ID No: g414810
US-10-883-805-16

  Query Match             91.0%;  Score 2505;  DB 17;  Length 486;
  Best Local Similarity   99.6%;  Pred. No. 6e-136;
  Matches  482;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         17 MMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGS 76
              |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGS 60

Qy         77 SISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGS 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGS 120

Qy        137 SSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSS 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSS 180

Qy        197 SSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVV 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVV 240

Qy        257 VDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKG 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKG 300

Qy        317 KIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGG 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGG 360

Qy        377 VQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSS 436
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        361 VQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGSPYHPCGSASQSPCSPPGTGSFSSSSS 420
```

```
Qy        437 SQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPC 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPC 480

Qy        497 RSIR 500
              ||||
Db        481 RSIR 484


RESULT 9
US-10-029-386-34046
; Sequence 34046, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34046
;   LENGTH: 219
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AP000510.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.21
;   OTHER INFORMATION: SWISSPROT HIT: Q15517, EVALUE 1.00e-24
US-10-029-386-34046

  Query Match             40.2%;  Score 1106;  DB 14;  Length 219;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-56;
  Matches  219;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         30 TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 60

Qy         90 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG 149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG 120

Qy        150 SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQ 209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQ 180

Qy        210 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSG 248
              ||||||||||||||||||||||||||||||||||||||||
Db        181 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSG 219


RESULT 10
US-10-316-175-71
; Sequence 71, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71
;   LENGTH: 146
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-71

  Query Match             26.1%;  Score 717;  DB 15;  Length 146;
  Best Local Similarity   100.0%;  Pred. No. 7.7e-34;
  Matches  146;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         55 CLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQ 114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQ 60

Qy        115 VSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGS 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGS 120

Qy        175 ALPTNDNSYRGILNPSQPGQSSSSSQ 200
              ||||||||||||||||||||||||||
Db        121 ALPTNDNSYRGILNPSQPGQSSSSSQ 146


RESULT 11
US-10-316-175-72
; Sequence 72, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 72
;   LENGTH: 111
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-72

  Query Match             19.4%;  Score 533;  DB 15;  Length 111;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-23;
  Matches  111;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         65 FSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSS 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 FSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSS 60

Qy        125 LQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSA 175
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSA 111


RESULT 12
US-10-316-175-73
; Sequence 73, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
```

```
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 73
;   LENGTH: 81
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-73

  Query Match             15.3%;  Score 421;  DB 15;  Length 81;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-17;
  Matches   81;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        375 GGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSS 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSS 60

Qy        435 SSSQSSGKIILQPCGSKSSSS 455
              |||||||||||||||||||||
Db         61 SSSQSSGKIILQPCGSKSSSS 81


RESULT 13
US-10-316-175-70
; Sequence 70, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 70
;   LENGTH: 58
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-70

  Query Match             11.3%;  Score 310;  DB 15;  Length 58;
  Best Local Similarity   100.0%;  Pred. No. 6.9e-11;
  Matches   58;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        472 DGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 DGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 58


RESULT 14
US-10-316-175-69
; Sequence 69, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 69
;   LENGTH: 55
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-69

  Query Match             10.8%;  Score 298;  DB 15;  Length 55;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-10;
  Matches   55;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPC 55
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPC 55


RESULT 15
US-09-840-746-20
; Sequence 20, Application US/09840746
; Publication No. US20030166501A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, Huei-Mei
;  APPLICANT:  Honchell, Cynthia D.
;  APPLICANT:  Tang, Y. Tom
;  TITLE OF INVENTION: Mucin-Related Tumor Marker
;  FILE REFERENCE: PC-0039 US
;  CURRENT APPLICATION NUMBER: US/09/840,746
;  CURRENT FILING DATE:  2001-04-23
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PERL Program
; SEQ ID NO 20
;   LENGTH: 528
;   TYPE: PRT
;   ORGANISM: Sus scrofa
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Genbank ID No. US20030166501A1 g915208
US-09-840-746-20

  Query Match             10.6%;  Score 292.5;  DB 10;  Length 528;
  Best Local Similarity   27.7%;  Pred. No. 6.9e-09;
  Matches  139;  Conservative   58;  Mismatches  240;  Indels   65;  Gaps   16;

Qy         54 PCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYS 113
              |  :    ::  :|   ||||   | |  |    ||||:  :|     ::  |  |    |
Db          6 PSSSSSSPTTSTTSVQSSSSSSVPIPSTTSVQPSSSGSAPTTSATSVQTSSSSSPPIS-S 64

Qy        114 QVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNG 173
              :|  : | ||:   | :|   |||| : :: : |  ||| ||:  |   |::| |  :
Db         65 TISVQTSSSSSVPTTSTTSVQPSSSSSAPTTRATSVQSSSSSSAPIS---STTSVQPSSS 121

Qy        174 SALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIV 233
              ::||         : :    |||||  :  ::| |  ||:  | || :    | |
Db        122 GSVPTT--------SATSVQSSSSSSAPTTSATSVQPSSSSSPPISSTVSVQPSSSSSAP 173

Qy        234 SHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAP--GVVQGPPCSNGGLP------------ 279
              : |   :  | | |       | |      | |         | |:  :|
Db        174 TTSATSVQPSSSSSPPISSTVSVQTSSSSSVPTTSTTSVQPSSSSSVPTTSATSVRSSSS 233

Qy        280 -GKPCPPITSVDKS-------YGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVK 331
               | |  |||  |          |   || | :|  |      :|    :  |
Db        234 SSTPIPSTTSVQPSSSSSAPTTSATSVQPSSSSSTPIPSTT----SVQPSS--SSSAPTT 287

Qy        332 GSPGV-PSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGG----GSTG 386
              :  | || :: ||||     |  |   | | :|  : ||  :|                |:
Db        288 SATSVQPSSSSSPPIS--STISVQPSSSSSSPTTSTTSVQPSSSGSAPTTSATSVQPSSS 345
```

```
Qy        387 SKGPCSPSSSRVPSSSSIS----------SSAGSPYHPCGSAS----QSPCSPPGTGSFS 432
              |  | | : |  ||||| |          ||:||   |  ||:     |  | | | : |
Db        346 SSPPISSTISVQPSSSSSSPTTSTTSVQPSSSGSA--PTTSATSVQPSSSSSVPTTSATS 403

Qy        433 SSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAG 492
                |||  ||  |         ||||  |  | :|:  |     :| |  :   :|  ||||
Db        404 VRSSSSSSTPIPTTTSVQPSSSSSVPTTSATSVQ-TSSSSSTPIPS-TTSVQPSSSSSAP 461

Qy        493 KIPCRSIRDILAQVKPLGPQLA 514
                |::   :   |:   ::
Db        462 TTSATSVQPSSSSSPPISSTIS 483
```

Search completed: August 2, 2005, 09:26:16
Job time : 165 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 2, 2005, 08:59:02 ; Search time 42 Seconds
(without alignments)
1211.873 Million cell updates/sec

Title: US-10-677-865-1
Perfect score: 2752
Sequence: 1 MGSSRAPWMGRVGGHGMMAL.....GPQLADPEVFLPQGELLDSP 529

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2583	93.9	502	2	A48679	differentiated ker
2	335	12.2	13288	2	T03099	mucin, submaxillar
3	330.5	12.0	1032	2	T34433	hypothetical prote
4	330.5	12.0	2232	2	T34434	hypothetical prote
5	328.5	11.9	1428	2	T08852	lustrin A - Califo
6	306.5	11.1	967	2	S66852	hypothetical prote
7	299.5	10.9	1585	2	T31611	hypothetical prote
8	292.5	10.6	528	2	I47141	gastric mucin (clo
9	289.5	10.5	1063	2	D86731	hypothetical prote
10	274.5	10.0	481	2	A35628	loricrin - mouse
11	272	9.9	1275	2	T33369	hypothetical prote
12	263	9.6	286	2	C61615	sericin MG-2 - gre
13	260	9.4	534	2	T39903	serine-rich protei
14	257	9.3	1459	2	T32271	hypothetical prote
15	254	9.2	1560	2	T02885	peroxisome prolife
16	253.5	9.2	1217	2	S52714	sericin1B - silkwo
17	250.5	9.1	536	2	T37544	hypothetical serin
18	248.5	9.0	383	2	T46707	proteophosphoglyca
19	248.5	9.0	3507	2	T34513	hypothetical prote
20	248	9.0	1367	1	S48478	glucan 1,4-alpha-g
21	244	8.9	627	2	A44112	spidroin 2, dragli
22	241.5	8.8	316	1	A38743	loricrin - human
23	241.5	8.8	786	2	T16509	hypothetical prote
24	240.5	8.7	805	2	T25795	hypothetical prote
25	238.5	8.7	406	2	S38170	SRP40 protein - ye
26	238	8.6	683	2	A82704	1,4-beta-cellobios
27	236.5	8.6	1464	2	S59856	collagen alpha 1(I
28	235.5	8.6	995	2	S50358	hypothetical prote
29	233	8.5	1896	1	RNFF2L	DNA-directed RNA p
30	231.5	8.4	1932	1	A28490	DNA-directed RNA p
31	230	8.4	1414	1	S23809	collagen alpha 2(I
32	230	8.4	1589	2	T42233	submaxillary mucin
33	229.5	8.3	539	2	T28770	hypothetical prote
34	229	8.3	725	2	A41258	a-agglutinin core
35	226.5	8.2	2468	2	A83412	hypothetical prote
36	224	8.1	1777	2	T34369	hypothetical prote
37	223.5	8.1	1306	2	S25370	MSB2 protein - yea
38	223	8.1	467	2	A27677	DNA-directed RNA p
39	223	8.1	4776	2	E95206	cell wall surface
40	221.5	8.0	888	2	T46726	secreted acid phos
41	221	8.0	1970	1	S21054	DNA-directed RNA p
42	221	8.0	1970	2	I38186	RNA polymerase II
43	219.5	8.0	1475	2	S42718	nuclear pore compl
44	219	8.0	937	2	S58135	hyphally regulated
45	218	7.9	1162	2	JH0557	exo-alpha-sialidas

ALIGNMENTS

RESULT 1
A48679
differentiated keratinocyte S protein precursor - human
C;Species: Homo sapiens (man)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C;Accession: A48679
R;Zhou, Y.; Chaplin, D.D.
Proc. Natl. Acad. Sci. U.S.A. 90, 9470-9474, 1993
A;Title: Identification in the HLA class I region of a gene expressed late in keratino
A;Reference number: A48679; MUID:94022396; PMID:8415725
A;Accession: A48679
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-502 <ZHO>
A;Cross-references: UNIPROT:O43509; GB:L20815
A;Note: sequence shown follows the authors' translation at positions 47 and 410
C;Genetics:
A;Map position: 6p21.3
C;Superfamily: loricrin

Query Match 93.9%; Score 2583; DB 2; Length 502;
Best Local Similarity 99.2%; Pred. No. 3.6e-115;
Matches 496; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              |||||||||||||||:|||||||||||||||||||||||||||| ||||||||||||
Db          1 MGSSRAPWMGRVGGHGMLALLLAGLLLPGTLAKSIGTFSDPCKDPTCITSPNDPCLTGKG 60

Qy         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180

Qy        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240

Qy        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300

Qy        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360

Qy        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||:| ||||||||||
Db        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGLPYHPCGSASQ 420
```

```
Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480

Qy        481 AGAKPCGSSSAGKIPCRSIR 500
              ||||||||||||||||||||
Db        481 AGAKPCGSSSAGKIPCRSIR 500
```

RESULT 2
T03099
mucin, submaxillary - pig
N;Alternate names: apomucin
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T03099; A40009; A28528; B29789
R;Eckhardt, A.E.; Timpte, C.S.; DeLuca, A.W.; Hill, R.L.
J. Biol. Chem. 272, 33204-33210, 1997
A;Title: The complete cDNA sequence and structural polymorphism of the polypeptide chain
A;Reference number: Z14839; MUID:98070526; PMID:9407109
A;Accession: T03099
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-13288 <ECK>
A;Cross-references: UNIPROT:O18758; EMBL:AF005273; NID:g2581863; PIDN:AAC62527.1; PID:g2
R;Eckhardt, A.E.; Timpte, C.S.; Abernethy, J.L.; Zhao, Y.; Hill, R.L.
J. Biol. Chem. 266, 9678-9686, 1991
A;Title: Porcine submaxillary mucin contains a cystine-rich, carboxyl-terminal domain in
A;Reference number: A40009; MUID:91236743; PMID:2033060
A;Accession: A40009
A;Molecule type: mRNA
A;Residues: 12139-12167,'T',12169-13288 <EC3>
A;Cross-references: GB:M61883; NID:g454837; PIDN:AAA30998.1; PID:g164374
R;Timpte, C.S.; Eckhardt, A.E.; Abernethy, J.L.; Hill, R.L.
J. Biol. Chem. 263, 1081-1088, 1988
A;Title: Porcine submaxillary gland apomucin contains tandemly repeated, identical seque
A;Reference number: A28528; MUID:88087170; PMID:2826455
A;Accession: A28528
A;Molecule type: mRNA
A;Residues: 12139-12167,'T',12169-12641 <TIM>
A;Cross-references: GB:M21174; GB:J03512; NID:g164321; PIDN:AAA30990.1; PID:g552360
A;Experimental source: submaxillary gland
R;Eckhardt, A.E.; Timpte, C.S.; Abernethy, J.L.; Toumadje, A.; Johnson Jr., W.C.; Hill,
J. Biol. Chem. 262, 11339-11344, 1987
A;Title: Structural properties of porcine submaxillary gland apomucin.
A;Reference number: A92606; MUID:87280230; PMID:3611111
A;Accession: B29789
A;Molecule type: protein
A;Residues: 1572-1607 <EC2>
C;Superfamily: pig submaxillary mucin
C;Keywords: tandem repeat

  Query Match   12.2%; Score 335; DB 2; Length 13288;
  Best Local Similarity 28.2%; Pred. No. 8.8e-08;
  Matches 177; Conservative   69; Mismatches 206; Indels 176; Gaps   31;

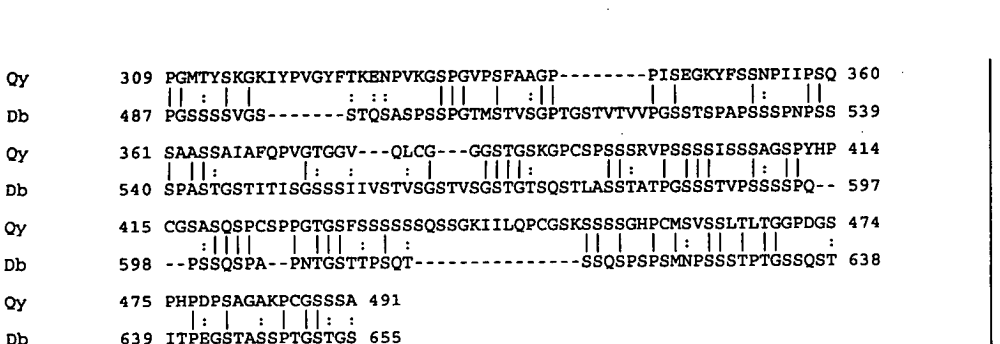
```
Qy          2 GSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKGD 61
              ||| || :   |          ||   ||| | |:|           :|    | : | |
Db       1576 GSSGAPAVSS-GASQAAGTSGAG---PGTTASSVG-----------VTETARPSVAGSGT 1620

Qy         62 S---SGFSSYSGSSS-----SGSSI---SSARSSGGGSSGS---SSGSSIAQGGSAGSFK 107
              :   || |  :||||     :|:||    ::| |  ||||:   |||:| | | |
Db       1621 TGTVSGASGSTGSSSGSPGATGASIGQPETSRISVAGSSGAPAVSSGASQAAGTSGAG-- 1678

Qy        108 PGTGYSQVSYS---------SGSGSSLQGASGSSQLGSSSSHSGSSGSHSG----SSSSH 154
              |||  | |  :          ||:  :: |||||:  ||||    |::|:  |    | |
Db       1679 PGTTASSVGVTETARPSVAGSGTTGTVSGASGST--GSSSGSPGATGASIGQPETSRISV 1736

Qy        155 SSSSSSFQFSSSSFQ-VGNGSALPTNDNSYRGILNPSQP-----------GQSSSSSQT 201
              : || :   || : |  |  | |    |  |:   ::|           | | |:  :
Db       1737 AGSSGAPAVSSGASQAAGTSGAGPGTTASSVGVTETARPSVAGSGTTGTVSGASGSTGSS 1796

Qy        202 SGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSH----------SGP-YIPSSHSVSGGQ 250
```

```
              ||  | | : :|  :| :| | :  || | ||           :||    ||  |:
Db       1797 SG--SPGATGASIGQPETSRISVAGSSGAPAVSSGASQAAGTSGAGPGTTASSVGVTETA 1854

Qy        251 RPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLV-- 308
              || |     |||  | | |   | |   | |     |: :       | |||  :  |
Db       1855 RPSVA-----GSGTTGTVSGASGSTGSSSGSPGATGASIGQPETSRISVAGSSGAPAVSS 1909

Qy        309 ------------PGMTYSKGKIYPVGYFTKENP----------VKGSPGVPSFAAGPP-- 344
                        || | |      ||      |          | |: |    ::| |
Db       1910 GASQAAGTSGAGPGTTASS-----VGVTETARPSVAGSGTTGTVSGASGSTGSSSGSPGA 1964

Qy        345 --ISEGKYFSSNPIIPSQSAAS--SAIAFQPVGTGGV-----------------QLCGGG 383
                 | |:  :|   :   | |   |: | |  || |                  : | |
Db       1965 TGASIGQPETSRISVAGSSGAPAVSSGASQAAGTSGAGPGTTASSVGVTETARPSVAGSG 2024

Qy        384 ST-------GSKGPCS-----------PSSSRVPSSSSISSSAGSPYHPCGSASQSPCS 424
              :|       || |  |           | :||:    |:: |:|:|    |::  :  |
Db       2025 TTGTVSGASGSTGSSSGSPGATGASIGQPETSRI----SVAGSSGAPAVSSGASQAAGTS 2080

Qy        425 PPGTGSFSSS----------SSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPD- 472
                | |: :||          : | ::|  :    ||  |||| |  : :|:   | |:
Db       2081 GAGPGTTASSVGVTETARPSVAGSGTTG-TVSGASGSTGSSSGSPGATGASI---GQPET 2136

Qy        473 -------GSPHPDPSAGA-KPCGSSSAG 492
                    |  |  |:|| :  |:| ||
Db       2137 SRISVAGSSGAPAVSSGASQAAGTSGAG 2164
```

RESULT 3
T34433
hypothetical protein K06A9.1a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T34433
R;Geisel, C.; Gattung, S.
submitted to the EMBL Data Library, December 1996
A;Description: The sequence of C. elegans cosmid K06A9.
A;Reference number: Z21525
A;Accession: T34433
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1032 <GEI>
A;Cross-references: UNIPROT:P91365; EMBL:U80846; PIDN:AAC70889.1; GSPDB:GN00028; CESP:K
A;Experimental source: strain Bristol N2; clone K06A9
C;Genetics:
A;Gene: CESP:K06A9.1a
A;Map position: X
A;Introns: 38/1; 75/3; 103/3; 132/2; 158/2; 222/1; 839/1; 849/1; 875/3; 940/2; 966/2; 9

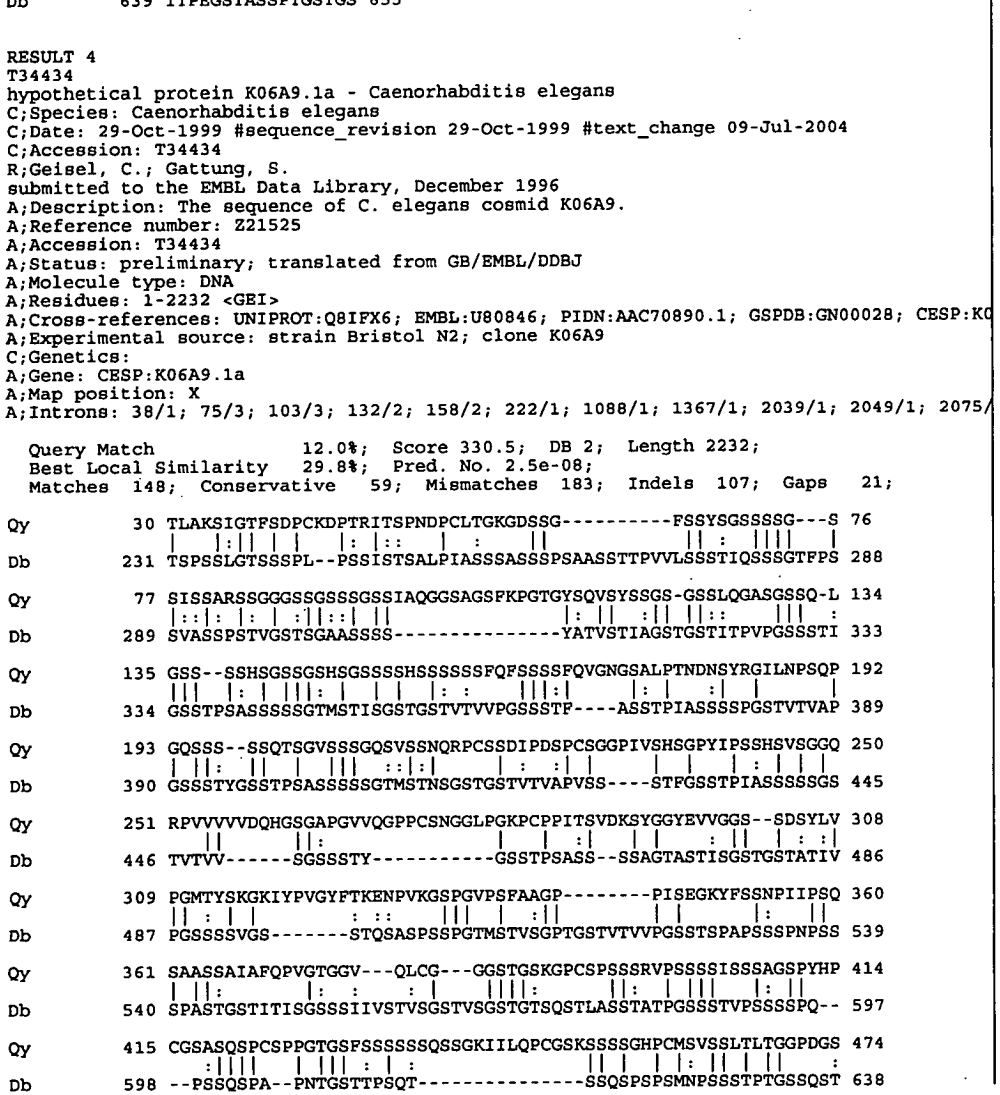
  Query Match   12.0%; Score 330.5; DB 2; Length 1032;
  Best Local Similarity 29.8%; Pred. No. 1.2e-08;
  Matches 148; Conservative   59; Mismatches 183; Indels 107; Gaps   21;

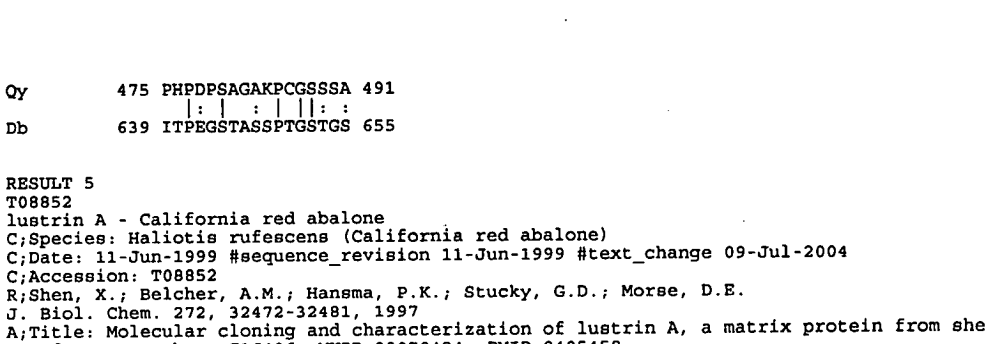
```
Qy         30 TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSG----------FSSYSGSSSSG---S 76
              |   |:|| | |   |: |::   |  :   ||           || :  |||   |
Db       231 TSPSSLGTSSSPL--PSSISTSALPIASSSASSSPSAASSTTPVVLSSSTIQSSSGTFPS 288

Qy         77 SISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGS-GSSLQGASGSSQ-L 134
              |::|: |: | :||::| ||                |: ||  :|| ||::    |||  :
Db       289 SVASSPSTVGSTSGAASSSS--------------YATVSTIAGSTGSTITPVPGSSSTI 333

Qy        135 GSS--SSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQP 192
              |||  |: | |||: |  | |  |: :    |||:|     |: |   :|  |       |
Db       334 GSSTPSASSSSSGTMSTISGSTGSTVTVVPGSSSTF----ASSTPIASSSSPGSTVTVAP 389

Qy        193 GQSSS--SSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQ 250
              | ||:  ||  |  |||  ::|:|      | :  :| |      |   |  | : | | |
Db       390 GSSSTYGSSTPSASSSSSGTMSTNSGSTGSTVTVAPVSS----STFGSSTPIASSSSSGS 445

Qy        251 RPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGS--SDSYLV 308
                  ||      ||:                |   |  :|   | |    : ||  | : :|
Db       446 TVTVV------SGSSSTY-----------GSSTPSASS--SSAGTASTISGSTGSTATIV 486
```



RESULT 4
T34434
hypothetical protein K06A9.1a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T34434
R;Geisel, C.; Gattung, S.
submitted to the EMBL Data Library, December 1996
A;Description: The sequence of C. elegans cosmid K06A9.
A;Reference number: Z21525
A;Accession: T34434
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2232 <GEI>
A;Cross-references: UNIPROT:Q8IFX6; EMBL:U80846; PIDN:AAC70890.1; GSPDB:GN00028; CESP:K0
A;Experimental source: strain Bristol N2; clone K06A9
C;Genetics:
A;Gene: CESP:K06A9.1a
A;Map position: X
A;Introns: 38/1; 75/3; 103/3; 132/2; 158/2; 222/1; 1088/1; 1367/1; 2039/1; 2049/1; 2075/

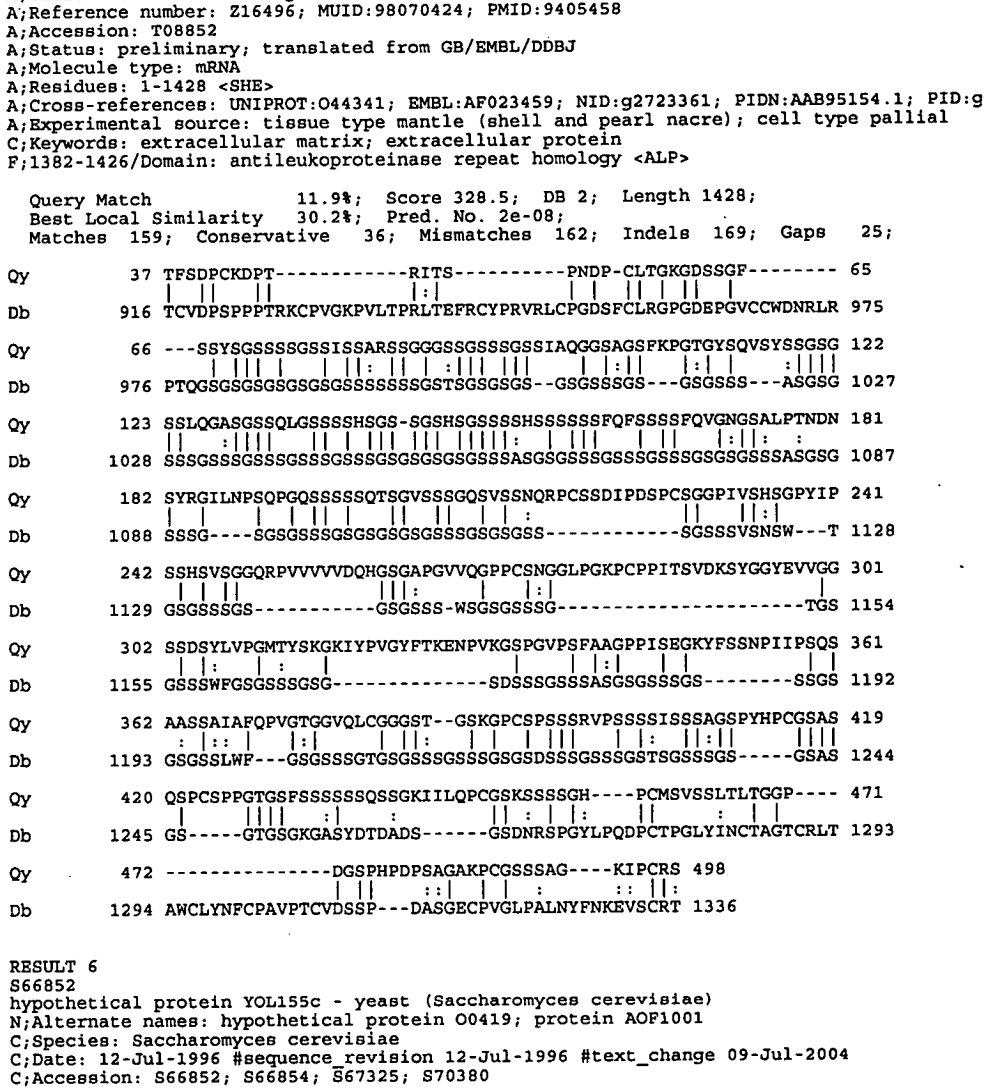
Query Match 12.0%; Score 330.5; DB 2; Length 2232;
Best Local Similarity 29.8%; Pred. No. 2.5e-08;
Matches 148; Conservative 59; Mismatches 183; Indels 107; Gaps 21;





RESULT 5
T08852
lustrin A - California red abalone
C;Species: Haliotis rufescens (California red abalone)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: T08852
R;Shen, X.; Belcher, A.M.; Hansma, P.K.; Stucky, G.D.; Morse, D.E.
J. Biol. Chem. 272, 32472-32481, 1997
A;Title: Molecular cloning and characterization of lustrin A, a matrix protein from she
A;Reference number: Z16496; MUID:98070424; PMID:9405458
A;Accession: T08852
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1428 <SHE>
A;Cross-references: UNIPROT:O44341; EMBL:AF023459; NID:g2723361; PIDN:AAB95154.1; PID:g
A;Experimental source: tissue type mantle (shell and pearl nacre); cell type pallial
C;Keywords: extracellular matrix; extracellular protein
F;1382-1426/Domain: antileukoproteinase repeat homology <ALP>

Query Match 11.9%; Score 328.5; DB 2; Length 1428;
Best Local Similarity 30.2%; Pred. No. 2e-08;
Matches 159; Conservative 36; Mismatches 162; Indels 169; Gaps 25;



RESULT 6
S66852
hypothetical protein YOL155c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein O0419; protein AOF1001
C;Species: Saccharomyces cerevisiae
C;Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
C;Accession: S66852; S66854; S67325; S70380

R;Arino, J.; Casamayor, A.; Gamo, F.J.; Gancedo, C.; Lafuente, M.J.; Aldea, M.; Casas, C
submitted to the Protein Sequence Database, July 1996
A;Reference number: S66814
A;Accession: S66852
A;Molecule type: DNA
A;Residues: 1-967 <ARI>
A;Cross-references: UNIPROT:Q08294; EMBL:Z74897; NID:g1420063; PID:e251930; PID:g1420064
A;Experimental source: strain S288C
R;Gaillon, L.; Dujon, B.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S66854
A;Accession: S66854
A;Molecule type: DNA
A;Residues: 1-967 <GAI>
A;Cross-references: EMBL:Z74897; NID:g1420063; PID:e251930; PID:g1420064; MIPS:YOL155c
A;Experimental source: strain S288C
R;Gamo, F.J.; Lafuente, M.J.; Casamayor, A.; Aldea, M.; Casas, C.; Ario, J.; Herrero, E.
submitted to the EMBL Data Library, July 1995
A;Description: Analysis of the DNA sequence of a 15500 bp fragment of the left arm of ch
pen reading frames.
A;Reference number: S67324
A;Accession: S67325
A;Molecule type: DNA
A;Residues: 1-164,'STSITSGSSSATESGSSVSGSTSATESGSSASGSS',166-186,'V',188-967 <GAM>
A;Cross-references: EMBL:X89715; NID:g1177620; PID:e190152; PID:g1177622
R;Gamo, F.J.; Lafuente, M.J.; Casamayor, A.; Arino, J.; Aldea, M.; Casas, C.; Herrero, E
Yeast 12, 709-714, 1996
A;Title: Analysis of the DNA sequence of a 15,500 bp fragment near the left telomere of
and two new open reading frames.
A;Reference number: S70379; MUID:96405919; PMID:8810044
A;Accession: S70380
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 55-164,'STSITSGSSSATESGSSVSGSTSATESGSSASGSS',166-186,'V',188-286;745-773;781
A;Cross-references: EMBL:X89715
C;Genetics:
A;Cross-references: SGD:S0005515
A;Map position: 15L
A;Note: YOL155c

Query Match 11.1%; Score 306.5; DB 2; Length 967;
Best Local Similarity 25.7%; Pred. No. 1.5e-07;
Matches 151; Conservative 73; Mismatches 210; Indels 153; Gaps 23;

```
Qy         45 PTRITSPNDPCLTGKGDS-SGFSSYSGSSSSGSSISSARSSGGGSSGSS----SGSSIA- 98
Db         83 PSTSSSEVSSSITSSGSSVSGSSSITSSGSSVSSSSSATESGSSASGSSSATESGSSVSG 142

Qy         99 -----QGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGS---SGSHSGS 150
Db        143 SSTSITSGSSSATESGSSVSGSTSATESGSSASGSSSATESGSSASGSSSATESGSSVSG 202

Qy        151 SSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTS------GV 204
Db        203 SSSATESGSSVSGSSSATESGSASSVPSSSGSV------TESGSSSSASESSITQSGTAS 256

Qy        205 SSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQ--RPVVVVVDQHGS 262
Db        257 GSSASSTSGSVTQSGSSVSGSSASSAPGISSSIPQSTSSASTASGSITSGTLSSITSSAS 316

Qy        263 GAPGVVQGPPCSNGGLPGKPCPPIT-------------SVDKSYGG-YEVVGG-----SS 303
Db        317 SATATASNSLSSSDGTIYLPSTTISGDITLTGSVIATEAVEVAAGGKLTLLDGDKYVFSA 376

Qy        304 DSYLVPGMTYSKGK-IYP-------------VGYFTKENPVKGSPGVPSFAAG------- 342
Db        377 DFIIHGGVFVEKSKPTYPGTEFDISGENFDVSGTFNAEEPAASSASAYSFTPGSFDNSGD 436

Qy        343 ----------------PPISEGKYFSSNPIIPSQSAA--------------------SS 365
Db        437 ISLSLSESTKGEVTFSPYSNSGAFSFSNAILNGGSVSGLQRRAESGSVNNGEINLENGST 496

Qy        366 AIAFQPV-GTGGVQLCGGG--------STGS----KG-------PCSPSSSRVP------ 399
```

```
Db        497 YVVVEPVSGSGTINIISGNLYLHYPDTFTGQTVVFKGEGVLAVDPTETNTTPIPVVGYTG 556

Qy        400 ----------SSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKI-----I 444
Db        557 ENQIAITADVTALSYDSATGVLTATQGNSQFS--FSIGTG-FSSSGFNVSEGTFAGAYAY 613

Qy        445 LQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSA 491
Db        614 YLNYGGVVASSATP----SSTSTTSGATNSTSGSTSFGASVTGSTAS 656
```

RESULT 7
T31611
hypothetical protein Y50E8A.g - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T31611
R;Steward, C.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z21047
A;Accession: T31611
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1585 <WIL>
A;Cross-references: EMBL:AL117200; NID:e1549770; PIDN:CAB55050.1; CESP:Y50E8A.g
A;Experimental source: clone Y50E8A
C;Genetics:
A;Gene: CESP:Y50E8A.g
A;Introns: 25/3; 60/1; 133/2; 217/3; 270/3; 337/2; 400/1; 746/2

Query Match 10.9%; Score 299.5; DB 2; Length 1585;
Best Local Similarity 24.5%; Pred. No. 5.1e-07;
Matches 155; Conservative 35; Mismatches 216; Indels 227; Gaps 23;

```
Qy          2 GSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSD--------PCKDPTRITSPND 53
Db        758 GSSGGGYASGGGGGG-----------GSSGGSGGRYSSAAAPPPPPPPPPAPAPAPSS 805

Qy         54 PCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQ-------------- 99
Db        806 GGYSGGSSGDSAAGGGGGSSRGYSEGSSSGGGGGSSGGYSGGSAAPPPPPPPPPAPAPA 865

Qy        100 ----------GGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSG------- 142
Db        866 PAPSSGGYSGGSSGGSAAGGGGGSSGGYSGGSSS---GGGG----GSSGGYSGGSAALPP 918

Qy        143 ---------------------SSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALP---- 177
Db        919 PPPPPPPPPPPPPAPAPAPAPSSGGYSGASSGGSAAGGG---GGSSGGYSGGSAAPPPPP 975

Qy        178 ------------------TNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCS 219
Db        976 PPPPPPPPPPPAPAPAPAPSSGGYSG-------GSSGGSAAGGGGGSSGGYSGGSAAPPP 1028

Qy        220 SDIPDSPCSGGPIVSHSGPYI--PSSHSVSGGQRPVVVVVDQHGSGAPGVVQG----PPC 273
Db       1029 PPPPPPPAPAPAPAPSSGGYSGGSSGGSAAGG-----------GGGSSGGYSGGSTAPPP 1077

Qy        274 SNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGS 333
Db       1078 PPPPPPPPPPAPAPAPAPSSGGYS--GGSSGGSAAGGGGGSSG-----GY-------SGG 1123

Qy        334 PGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKG---- 389
Db       1124 SAAPPPPPPPP----------PPAPAPAPAPSSGGYSGGSSGGSAAGGGG--GSSGGYSG 1171

Qy        390 --------------PCSPSSSRVPSSSSIS--SSAGS----------------------P 411
Db       1172 GSAAPPPPPPPPPPPPAPAPAPAPSSGGYSGGSSGGSAAGGGGSSGGYSGGSAAPPPPP 1231

Qy        412 YHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGG- 470
```

```
Db       1232 PPPPPPPAPAPAPAPSSGGYSGGSSGGSAAG------GGGGSSGGY----------TGGS 1275

Qy        471 --------------PDGSPHPDPSAGAKPCGSS 489
                           |  :| | ||:|    |||
Db       1276 AAPPPPPPPPPPPPPAPAPAPAPSSGGYSGGSS 1308


RESULT 8
I47141
gastric mucin (clone PGM-2A) - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 09-Jul-2004
C;Accession: I47141; S55315
R;Turner, B.S.; Bhaskar, K.R.; Hadzopoulou-Cladaras, M.; Specian, R.D.; LaMont, J.T.
Gastroenterology 106, 200, 1994
A;Title: Pig gastric mucin: isolation and characterization of a cDNA clone with a novel
A;Reference number: I47141; MUID:94102478; PMID:7506218
A;Accession: I47141
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-528 <TUR>
A;Cross-references: UNIPROT:Q29071; EMBL:U10281; NID:g915205; PIDN:AAC48526.1; PID:g9152
R;Turner, B.S.; Bhaskar, K.R.; Hadzopoulou-Cladaras, M.; Specian, R.D.; LaMont, J.T.
Biochem. J. 308, 89-96, 1995
A;Title: Isolation and characterization of cDNA clones encoding pig gastric mucin.
A;Reference number: S55315; MUID:95275264; PMID:7755593
A;Accession: S55315
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-528 <TU2>
A;Cross-references: GB:U10281; NID:g915205; PIDN:AAC48526.1; PID:g915208
C;Superfamily: pig submaxillary mucin

  Query Match            10.6%;  Score 292.5;  DB 2;  Length 528;
  Best Local Similarity   27.7%;  Pred. No. 3.7e-07;
  Matches  139;  Conservative   58;  Mismatches  240;  Indels   65;  Gaps   16;

Qy         54 PCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYS 113
              |  :    ::  :|   ||||   | |  |    ||||:  :|     ::  |  |    |
Db          6 PSSSSSSPTTSTTSVQSSSSSSVPIPSTTSVQPSSSGSAPTTSATSVQTSSSSSPPIS-S 64

Qy        114 QVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNG 173
               :|  : | ||:   | :|   |||| : :: : |  ||| ||:  |   |::| |  :
Db         65 TISVQTSSSSSVPTTSTTSVQPSSSSSAPTTRATSVQSSSSSSAPIS---STTSVQPSSS 121

Qy        174 SALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIV 233
               ::||          : :    |||||  :  ::| |  ||:  | || :    | |
Db        122 GSVPTT--------SATSVQSSSSSSAPTTSATSVQPSSSSSPPISSTVSVQPSSSSSAP 173

Qy        234 SHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAP--GVVQGPPCSNGGLP------------ 279
              : |   :  | | |     | |     | |        | |:  :|
Db        174 TTSATSVQPSSSSSPPISSTVSVQTSSSSSVPTTSTTSVQPSSSSSVPTTSATSVRSSSS 233

Qy        280 -GKPCPPITSVDKS-------YGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVK 331
                | |  |||  |          |   || |  :|  |     : |   :   |
Db        234 SSTPIPSTTSVQPSSSSSAPTTSATSVQPSSSSSTPIPSTT----SVQPSS--SSSAPTT 287

Qy        332 GSPGV-PSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGG----GSTG 386
               :  | || :: ||||     |  |   | |  :|  : ||  :|             |:
Db        288 SATSVQPSSSSSPPIS--STISVQPSSSSSSPTTSTTSVQPSSSGSAPTTSATSVQPSSS 345

Qy        387 SKGPCSPSSSRVPSSSSIS----------SSAGSPYHPCGSAS----QSPCSPPGTGSFS 432
              |  | | : |  ||||| |          ||:||   |  ||:     |  | | | : |
Db        346 SSPPISSTISVQPSSSSSSPTTSTTSVQPSSSGSA--PTTSATSVQPSSSSSVPTTSATS 403

Qy        433 SSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAG 492
                |||  ||  |         ||||  |  | :|:  |     :| |  :   :|  ||||
Db        404 VRSSSSSSTPIPTTTSVQPSSSSSVPTTSATSVQ-TSSSSSTPIPS-TTSVQPSSSSSAP 461

Qy        493 KIPCRSIRDILAQVKPLGPQLA 514
                  |::   :   |:   ::
Db        462 TTSATSVQPSSSSSPPISSTIS 483


RESULT 9
D86731
hypothetical protein yihD [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: D86731
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrl
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: D86731
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1063 <STO>
A;Cross-references: UNIPROT:Q9CH86; GB:AE005176; PID:g12723779; PIDN:AAK04950.1; GSPDB:
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: yihD

  Query Match            10.5%;  Score 289.5;  DB 2;  Length 1063;
  Best Local Similarity   28.5%;  Pred. No. 1e-06;
  Matches  134;  Conservative   52;  Mismatches  169;  Indels  116;  Gaps   15;

Qy         27 LPGTLAKSIGTFSDPCKDPTRITSPNDPCLTG-------KGDSSGFSSYSGSSSSGS--S 77
              |||  |   |  :   : |  | :  :  | |              | | |: : :||: |  |
Db        556 LPGLGAIENGLSNTISQIPNNINNFVNNALNGITTIINSLTPSVGASTVNPNSSANSSQS 615

Qy         78 ISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGS- 136
              ||| ||   :| |:|  |:::   |: | :  |  |  | :| | || :|:| ||   |
Db        616 SSSASSSSSAASSSTSSSNVSSNTSSNSSEANTS-SSTSNASSSSSSSEGSSASSSNSSE 674

Qy        137 ---------SSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGIL 187
                       || | |:| :| |||:  ||:||   | ||  | : |::  :: :|  |:
Db        675 SSVASSSSVDSSQSSSAGVNSSSSSAEGSSASSSNSSESS--VASSSSVDSSQSSSAGV- 731

Qy        188 NPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVS 247
                  |||||::|  |||  |  |                    : || ||
Db        732 -----NGSSSSSESSSASSSNSSEGS-------------------------VASSSSVD 760

Qy        248 GGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYL 307
                |           | : ||      |:    |        | : |      || | |
Db        761 SSQ----------SSSAGV----NSSSSSAEGSSASSSNSSESSIASSSSVGSSQSS-- 803

Qy        308 VPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAI 367
                                    | || | ::    :||   ||:    |  |:||::
Db        804 ------------------------STGVSSSSSS---AEGSSASSSNSSESSVASSSSV 835

Qy        368 AFQPVGTGGVQLCGGGSTGSKGPCS-------PSSSRVPSSSSISSSAGSPYHPCGSASQ 420
                  : ||      | ||     |          ||| | || | ||: ||      |:
Db        836 DSSQSSSAGVNSSSSSSEGSSASSSNFSESSIGSSSSVDSSQSSSSTEGSSASSLSSSES 895

Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMS---VSSLTLT 468
              |  |  || | |:|:: |:|         | | :| |  :|    || :||
Db        896 SMTSSGGT-SDSNSNTGQTS------TTSSSSKNSSHSSLSSNESSSASLT 939


RESULT 10
A35628
loricrin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 21-Sep-1990 #sequence_revision 21-Sep-1990 #text_change 09-Jul-2004
C;Accession: A35628
R;Mehrel, T.; Hohl, D.; Rothnagel, J.A.; Longley, M.A.; Bundman, D.; Cheng, C.; Lichti
Cell 61, 1103-1112, 1990
A;Title: Identification of a major keratinocyte cell envelope protein, loricrin.
A;Reference number: A35628; MUID:90275605; PMID:2190691
A;Accession: A35628
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-481 <MEH>
A;Cross-references: UNIPROT:P18165; GB:M34398; NID:g198870; PIDN:AAA39444.1; PID:g1988
C;Superfamily: loricrin
```

```
C;Keywords: cornified cell envelope; epidermis

  Query Match             10.0%;  Score 274.5;  DB 2;  Length 481;
  Best Local Similarity   26.6%;  Pred. No. 2.4e-06;
  Matches  120;  Conservative   20;  Mismatches  164;  Indels  147;  Gaps   17;

Qy         58 GKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSG----------SSIAQGGSAGSFK 107
Db        136 GGGGSGGGVKYSGGGGGGGSSCGGGSSGGGGGGSSCGGGSGGGGSYCGGSSGGGSSGGCG 195

Qy        108 PGTGYSQVSYSSGSGSSLQG----ASGSSQLGSS--SSHSGSSGSHSGSSSSHSSSSSSF 161
Db        196 GGSGGGK--YSGGGGGSSCGGGYSGGGGSSGGSSCGGGYSGGGGSSCGGGGGYSGGGGTS 253

Qy        162 QFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSD 221
Db        254 CGGGSSGGGGGGSSQQYQCQSYGG-------GSSGGSSCGGGYSGGG------------- 293

Qy        222 IPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGK 281
Db        294 --GSSCGGG----YSG----------GG-----------GSSCGGGSSGGGSSCGGSGG- 325

Qy        282 PCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAA 341
Db        326 ------------GGYSGGGG-----------------------------GSCGGGSSGG 343

Qy        342 GPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSS 401
Db        344 G-----GGYYSSQQ--TSQTSCAPQQSYGGGSSGGGGSCGGGSSGGGG----------GG 386

Qy        402 SSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMS 461
Db        387 GCYSSGGGGSSGGCGGGYSG-----GGGGCGGGSSGGSGG-----GCGGGSSGG------ 430

Qy        462 VSSLTLTGGPDGSPHPDPSAGAKPCGSSSAG 492
Db        431 ------SGGGCGGGYSGGGGGGSSCGGGSSG 455


RESULT 11
T33369
hypothetical protein H02F09.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T33369
R;Geisel, C.; Harmon, G.
submitted to the EMBL Data Library, July 1998
A;Description: The sequence of C. elegans cosmid H02F09.
A;Reference number: Z21330
A;Accession: T33369
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1275 <GEI>
A;Cross-references: UNIPROT:O76602; EMBL:AF077538; PIDN:AAC64622.1; GSPDB:GN00028; CESP:
A;Experimental source: strain Bristol N2; clone H02F09
C;Genetics:
A;Gene: CESP:H02F09.3
A;Map position: X
A;Introns: 42/1; 78/3; 106/3; 135/2; 161/2; 224/1
C;Superfamily: yeast glucan 1,4-alpha-glucosidase homolog; glucan 1,4-alpha-glucosidase

  Query Match             9.9%;  Score 272;  DB 2;  Length 1275;
  Best Local Similarity   24.8%;  Pred. No. 8.2e-06;
  Matches  145;  Conservative   66;  Mismatches  244;  Indels  130;  Gaps   19;

Qy         28 PGTLAKSIGTFSDPCKDPTR---ITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSS 84
Db        572 PTSAGSSISTVSTVSSQPSTYIPVSSASSIYSTLSGSTGSTASPGTTESSGSSTSGPSTI 631

Qy         85 GGGSSGSSSSGSSIAQ-----GGSAGSFKPGTGYSQVSYSSG-SGSSLQGASGSSQLGS-- 136
Db        632 SGSSASTVTGSTVTEASTISGSTESSTIPGSTESTVSEASTVSGSSVSTVSGSTESTSAG 691

Qy        137 SSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSS 196
Db        692 ASTVSGSTGSTVSDSSTISDSTGSTNAPGSTESTVTGSSVSTVSGS-TGSTGPSTMSAST 750

Qy        197 SSSQTSG-----------VSSSGQSVSSNQRPCSSDIPD-----------SPCSG--GPI 232
Db        751 GSTNTPGSTESTITDGSTVSGSTGSTGSTNNPGSTDSSTTGISTVSGSSLSTISGSTGST 810

Qy        233 VSHSGPYIPSSHSVS--GGQRPVV-------------VVVDQHGSGAPGVVQGPPCSNGG 277
Db        811 VSGSSDMTVSTGSTSSPGSTESTVSGASTMSPSTGSSVETSTSGSSVSTVSQSTSSSTTG 870

Qy        278 ----------------------------------------------------------LPGKPC 283
Db        871 QSTVSESSVSTVSSESTISQSTGSTTTGESTVFGSTGSTATGSSTMSASTGSTDTPGSTE 930

Qy        284 PPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGY-----FTKENPVKGSPGVPS 338
Db        931 STITG--STVTGESTVSGSTGSTITEGSTISESTMTTVGVSTGSTITGESTVSGS--TRS 986

Qy        339 FAAGPPISEGKYFS-------SNPIIPSQSAAS--SAIAFQPVGTGGVQLCGGGSTGSKG 389
Db        987 TVTGESTVSGSTESTVSGSTESTPTVPSTVSGSTGSTVTGESTVSGSTASTSSGSTGSST 1046

Qy        390 PCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCG 449
Db       1047 EAGSTVSG-SSASTVTSSTGS-----STSGESTVSGSTVSTVSGSTGSTITGESTVSG-S 1099

Qy        450 SKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKI 494
Db       1100 TESTVTAESTVSGSSVSTVSGNTGSTITGESTVSGSTGSTGESTI 1144


RESULT 12
C61615
sericin MG-2 - greater wax moth (fragments)
N;Alternate names: silk protein MG-2
C;Species: Galleria mellonella (greater wax moth)
C;Date: 18-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 28-Jul-1995
C;Accession: C61615
R;Zurovec, M.; Sehnal, F.; Scheller, K.; Kumaran, A.K.
Insect Biochem. Mol. Biol. 22, 55-67, 1992
A;Title: Silk gland specific cDNAs from Galleria mellonella L.
A;Reference number: A61615
A;Accession: C61615
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-286 <ZUR>
C;Keywords: tandem repeat

  Query Match             9.6%;  Score 263;  DB 2;  Length 286;
  Best Local Similarity   28.6%;  Pred. No. 5e-06;
  Matches  114;  Conservative   41;  Mismatches  108;  Indels  136;  Gaps   12;

Qy         57 TGKGDSSGFSS---YSGSSSSGSSISSARSSGGGSSGSSSGSSSIAQGGSAGSFKPGTGYS 113
Db          5 SSSNNSSGSSSSNNSSGSSSNNSSGSSSNNSSGSSSNNSSGSSSSNNSSGSSSNNSSGSS 64

Qy        114 QVSYS---SGSGSSLQGASGSSQL----GSSSSH--SGSSGSH--SGSSSSHSSSSSSFQ 162
Db         65 SNSSSTNNSSGSSSTNNSSGSSSTNNSSGSSSTNNASGSSSSNNTSGSSRNNSSGSSSSN 124

Qy        163 FSSSSFQVGNGS-ALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSD 221
Db        125 NSSSGSSAGNSSGSSSSNSSGSSGIGSLSRNSSSSSSSQAAG-SSSSRRVSADGSSSSSS 183

Qy        222 IPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGK 281
Db        184 ASNSAA--------------AQNSASSSE-----TINADGS------------------- 205

Qy        282 PCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAA 341
Db        206 ------------------------------------------EN------------- 207
```

```
Qy        342 GPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSS 401
                          :|::||: | |    |    :   ||  |    | :|::  ::
Db        208 -----------------ESSSSSSSAAQNSATRSQVINADGSQSSSSSSSSASNQASAT 249

Qy        402 SSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSS 440
              || | ||       || |:|  |   : | || ||| ||
Db        250 SSSSVSAD------GSESESSSSSSSSSSSSESSSSSS 282


RESULT 13
T39903
serine-rich protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T39903
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.
submitted to the EMBL Data Library, November 1998
A;Reference number: Z21889
A;Accession: T39903
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-534 <LYN>
A;Cross-references: UNIPROT:O94317; EMBL:AL033534; PIDN:CAA22127.1; GSPDB:GN00067; SPDB:
A;Experimental source: strain 972h-; cosmid c215
C;Genetics:
A;Gene: SPDB:SPBC215.13
A;Map position: 2
C;Superfamily: pig submaxillary mucin

  Query Match             9.4%;  Score 260;  DB 2;  Length 534;
  Best Local Similarity   28.1%;  Pred. No. 1.3e-05;
  Matches  128;  Conservative   45;  Mismatches  132;  Indels  150;  Gaps   15;

Qy         19 ALLLAGLLLPGTLAKSIGTFSDPCKDPT-----RITSPNDPCLTGKGDSSGFSSYSGSSS 73
              |:| :   :  ::: |  : |||     |        || : | ::    || ||| : :|:
Db        147 AILPSSTSVEVSISSSSLSSSDPLTSSTFSSLSSSTSSSQPSVSST-SSSTFSSAAPTST 205

Qy         74 SGSSISSARSSGGGSSGSSSGSSIAQGGS-AGSFKPGTGYSQVSYSSGSGSSLQGASGSS 132
              | | :||:      || ||| ||     | : |  | |     | || : ||| :| ||
Db        206 SSSYLSSSSVVSSSSSPSSSSSSTLTSSSLSTSSIPSTS----SSSSSTSSSLSSSSSSS 261

Qy        133 QLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQP 192
              |||| |    | | |||| :|:||:   ||||      |: ||: :|
Db        262 TASSSSSSSSIISSSSSSSSSPTSTSSTISSSSSS------SSSPTSTSS---------- 305

Qy        193 GQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRP 252
              ||||| :|  ||   ::||:    ||    || |    :| |    ||| | |
Db        306 TISSSSSSSSSFSS---TLSSSSMSSSSSFSSSPTSSSSTISSSSS-SPSSSSFSS---- 357

Qy        253 VVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMT 312
                                                 ||  |          ||.|:
Db        358 --------------------------------TTSSSK----------SSSSF------ 368

Qy        313 YSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPV 372
                                                 |: :  | | :|| :
Db        369 ---------------------------------------SSTVSSSSSTSSSTL----- 383

Qy        373 GTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFS 432
                  | |    | ||||  |||| ||| |  |  |  |:|:|  :|  :  :
Db        384 -----------TSS----SSSSSRPASSSSHSSSLSS--HKSSSSSKSSSAPVSSAFYH 425

Qy        433 SSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTL 467
              :|:||:||        | || |  | :: || :|
Db        426 NSTSSRSSSH---SSSHSLSSLSSKPILTASSSSL 457


RESULT 14
T32271
hypothetical protein ZC178.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T32271
R;Murray, J.; Wohldmann, P.; Beck, C.
submitted to the EMBL Data Library, September 1997
A;Description: The sequence of C. elegans cosmid ZC178.
A;Reference number: Z21143
A;Accession: T32271
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1459 <MUR>
A;Cross-references: UNIPROT:O17084; EMBL:AF024496; PIDN:AAB70340.1; GSPDB:GN00023; CES
A;Experimental source: strain Bristol N2; clone ZC178
C;Genetics:
A;Gene: CESP:ZC178.2
A;Map position: 5
A;Introns: 1099/2

  Query Match             9.3%;  Score 257;  DB 2;  Length 1459;
  Best Local Similarity   22.4%;  Pred. No. 4.8e-05;
  Matches  137;  Conservative   86;  Mismatches  224;  Indels  164;  Gaps   22;

Qy         28 PGTLAKSIGTFSDPCKDPTR-------ITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISS 80
              | : | :  :   ||   |:        :| |||    |        |: : |:|| :  :|
Db        202 PSSTATNSASSETPCNSETQTSDGTSTMTVPNDSTTAGPS-----STATNSASSETPCNS 256

Qy         81 ARSSGGGSS-------GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQ 133
                 :  |:|         :::| |      || | |    :|  :|   |::   : |:
Db        257 ETQTSDGTSTMTVPNDSTTAGPSSTATNSASSETPCNSETQ---TSDGTSTMTVPNDSTT 313

Qy        134 LGSSSSHSGSSGSH----------------------SGSSSSHSSSSSSFQFSSSSFQ 169
              | ||: : |: |                      :| ||: ::|:||    :|  |
Db        314 AGPSSTATNSASSETPCNSETQTSDGTSTMTVPNDSTTAGPSSTATNSASSETPCNSETQ 373

Qy        170 VGNGSALPT--NDNSYRGILNPSQPGQSSSSSQ----------------------TSGV 204
                :|::  |  ||::  |   ||    :|:||:                      |:|
Db        374 TSDGTSTMTVPNDSTTAG---PSSTATNSASSETPCNSETQTSDGTSTMTVPNDSTTAGP 430

Qy        205 SSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHS----GPYIPSSHSVSG------------ 248
              ||:  : :|:: ||:|:   |  :    ||      ||   :::| |
Db        431 SSTVTNSASSETPCNSETQTSDGTSTMTVSSDSTTPGPSSTATNSASSETPCNSETQTSD 490

Qy        249 GQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLV 308
              |   : |  |  :|   |     :|     ||    |  ::  |   :  ||||
Db        491 GTFTMTVSSDSTTAGPSSTV-----TNSASSETPC---NSETQTSDGTSTMTVSSDS-TT 541

Qy        309 PG----MTYSKGKIYPVGYFTKENPVKGSPGVP--SFAAGPPISEGKYFSSNPIIPSQSA 362
              ||    | |      |    |: :    :  ||  |  |||  :     ||     |::
Db        542 PGPSSTATNSASSETPCNSETQTSDGTSTMTVPNDSTTAGPSSTATNSASSETPCNSETQ 601

Qy        363 ASSAIAFQPV----GTGGVQLCGGGSTGSKGPCS--------------PSSSRVPSSSSI 404
              |   :  |      | |        |  |: ||:                | |    ||
Db        602 TSDGTSTMTVPNDSTTAGPSSTVTNSASSETPCNSETQTSDGTSTMTVSSDSTTAGPSST 661

Qy        405 SSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIIL--------QPCGSKS-SSS 455
              :::: |   || | :|   :  || : : || | ::|              || |:: :|
Db        662 ATNSASSETPCNSETQ---TSDGTSTMTVSSDSTTAGPSSTATNSASSETPCNSETQTSD 718

Qy        456 GHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGS-------------------------- 488
              |   |:||| : | ||  |   : ::   || |
Db        719 GTSTMTVSSDSTTAGP-SSTATNSASSETPCNSETQTSDGTFTMTVPNDSTTAGPSSTVT 777

Qy        489 -SSAGKIPCRS 498
               |:: : || |
Db        778 NSASSETPCNS 788


RESULT 15
T02885
peroxisome proliferator-activated receptor gamma binding protein PBP165 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T02885
R;Zhu, Y.; Qi, C.; Jain, S.; Rao, M.S.; Reddy, J.K.
J. Biol. Chem. 272, 25500-25506, 1997
A;Title: Isolation and characterization of PBP, a protein that interacts with peroxiso
```

A;Reference number: Z14760; MUID:97467333; PMID:9325263
A;Accession: T02885
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1560 <ZHU>
A;Cross-references: UNIPROT:O88323; EMBL:AF000294; NID:g3411010; PIDN:AAC31118.1; PID:g3

Query Match 9.2%; Score 254; DB 2; Length 1560;
Best Local Similarity 27.9%; Pred. No. 7.1e-05;
Matches 147; Conservative 35; Mismatches 179; Indels 166; Gaps 24;

```
Qy         25 LLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSS 84
              ||  | | |         |   |:   |     | | |||  |  || |     |  |:
Db        942 LLTTGELGKE--------KTQKRVKEGN-----GTGASSG--SGPGSDSKPGKRSRTPSN 986

Qy         85 GGGS--------SGSSSGSSIAQGGSAGSFKP--GTGYSQVSYSSGSGSSLQGAS----- 129
               | |           : | | :   |   | |   || |:   |||   :  | :
Db        987 DGKSKDKPPKRKKADTEGKSPSHSSSNRPFTPPTSTGGSKSPGSSGRSQTPPGVATPPIP 1046

Qy        130 --------GSSQLGSSSSHS--GSSGSHSGSSS----SHSSSSSSFQFSSSSFQVGNGSA 175
                     |:  :|  ||||   |||| | | |    ||||||||   :|:| :|
Db       1047 KITIQIPKGTVMVGKPSSHSQYTSSGSVSSSGSKSHHSHSSSSSSLASASTSGKV----- 1101

Qy        176 LPTNDNSYRGILNPSQPGQSSSSSQTSG-----VSSSGQSVSSNQRPCSSDIPDSPCSGG 230
                        |   : ||||: ||       ||| | | |    ||      | |
Db       1102 ------------KSSKSEGSSSSKLSGSMYASQGSSGSSQSKN----SSQTGGKPGS-S 1143

Qy        231 PIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGL-PGKPCPPITSV 289
              ||  |      || | |   :|          | |  :  |  |   : |     ||
Db       1144 PITKHG----LSSGSSSTKMKP---------QGKPSSLMNPSISKPNISPSHSRPP---- 1186

Qy        290 DKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGK 349
                    | ||    |                      || |:|  ||  |  ||| |
Db       1187 ----------GGSDKLASP----------------MKPVPGTP--PSSKAKSPISSG- 1215

Qy        350 YFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSP--SSSRVPSSSSISSS 407
               ||   :   |::|         |     |  |:||   :|  |:|  ||||| |||
Db       1216 --SSGSHVSGTSSSS----------GMKSSSGSASSGSVSQKTPPASNSCTPSSSSFSSS 1263

Qy        408 AGSPYHPCGSASQSPCSPPGTGSFSSSS--------------------SSQSSGKIILQP 447
                     ||:  |  :  |:    | |                    |   |   |
Db       1264 --------GSSMSSSQNQHGSSKGKSPSRNKKPSLTAVIDKLKHGVVTSGPGSEDPIDSQ 1315

Qy        448 CGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKI 494
               |: ::|| ||   :||   | | :     : |    |   |:| | :
Db       1316 MGASTNSSNHP---MSSKHNTSGGEFQSKREKSDKDKSKVSASGGSV 1359
```

Search completed: August 2, 2005, 09:13:11
Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 2, 2005, 08:43:06 ; Search time 172 Seconds
(without alignments)
1574.941 Million cell updates/sec

Title: US-10-677-865-1
Perfect score: 2752
Sequence: 1 MGSSRAPWMGRVGGHGMMAL.....GPQLADPEVFLPQGELLDSP 529

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2752	100.0	529	2	O43509	O43509 homo sapien
2	2749	99.9	529	2	Q8IZU4	Q8izu4 homo sapien
3	2731	99.2	529	2	Q7YR44	Q7yr44 pan troglod
4	2719.5	98.8	528	2	Q8IZU5	Q8izu5 homo sapien
5	2719.5	98.8	528	2	Q8IZU6	Q8izu6 homo sapien
6	2715.5	98.7	528	2	Q8N5P3	Q8n5p3 homo sapien
7	2657	96.5	513	2	Q95IF9	Q95if9 homo sapien
8	2596	94.3	500	2	Q9NP52	Q9np52 homo sapien
9	2593	94.2	500	2	Q9NPG5	Q9npg5 homo sapien
10	2590	94.1	500	2	Q9NRH8	Q9nrh8 homo sapien
11	2586	94.0	500	2	Q9NRH6	Q9nrh6 homo sapien
12	2586	94.0	500	2	Q9NRH7	Q9nrh7 homo sapien
13	2576	93.6	500	2	Q9NPE0	Q9npe0 homo sapien
14	2571.5	93.4	499	2	Q9NRH5	Q9nrh5 homo sapien
15	2569.5	93.4	499	2	Q9NRH4	Q9nrh4 homo sapien
16	2505	91.0	486	1	CDSN_HUMAN	Q15517 homo sapien
17	2214	80.5	424	2	Q9UBH8	Q9ubh8 homo sapien
18	2211	80.3	424	2	Q7LA70	Q7la70 homo sapien
19	2208	80.2	424	2	Q9UIP0	Q9uip0 homo sapien
20	2204	80.1	424	2	Q7LA71	Q7la71 homo sapien
21	2199	79.9	424	2	Q9UIN7	Q9uin7 homo sapien
22	2187.5	79.5	423	2	Q9UIN6	Q9uin6 homo sapien
23	1752	63.7	561	2	Q7TPC1	Q7tpc1 mus musculu
24	974	35.4	225	1	CDSN_PIG	O19084 sus scrofa
25	708	25.7	262	2	Q767M8	Q767m8 sus scrofa
26	704	25.6	133	2	Q7Z559	Q7z559 homo sapien
27	703	25.5	133	2	Q7Z560	Q7z560 homo sapien
28	370.5	13.5	1325	2	Q9BKV7	Q9bkv7 leishmania
29	355	12.9	669	2	Q7UE67	Q7ue67 rhodopirell
30	345	12.5	374	2	Q75JC9	Q75jc9 dictyosteli
31	340.5	12.4	950	2	Q9VE45	Q9ve45 drosophila
32	338	12.3	1416	2	Q6VBJ3	Q6vbj3 candida gla
33	335.5	12.2	1218	2	Q6VBJ4	Q6vbj4 candida gla
34	335	12.2	13288	2	O18758	O18758 sus scrofa
35	330.5	12.0	825	2	Q7YZW4	Q7yzw4 caenorhabdi
36	330.5	12.0	1032	2	P91365	P91365 caenorhabdi
37	330.5	12.0	2232	2	Q8IFX6	Q8ifx6 caenorhabdi
38	328.5	11.9	1428	2	O44341	O44341 haliotis ru
39	328	11.9	774	2	Q868B4	Q868b4 caenorhabdi
40	328	11.9	1035	2	Q76C74	Q76c74 saccharomyc
41	328	11.9	1713	2	Q8TGE1	Q8tge1 saccharomyc
42	320	11.6	2112	2	Q9VEL9	Q9vel9 drosophila
43	318	11.6	1553	2	Q6FPR0	Q6fpr0 candida gla
44	314.5	11.4	2850	1	HORN_HUMAN	Q86yz3 homo sapien
45	309.5	11.2	1376	2	Q7S5H8	Q7s5h8 neurospora

ALIGNMENTS

RESULT 1
O43509
ID O43509 PRELIMINARY; PRT; 529 AA.
AC O43509;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Corneodesmosin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Epidermis;
RX MEDLINE=98380490; PubMed=9712893; DOI=10.1074/jbc.273.35.22640;
RA Guerrin M., Simon M., Montezin M., Haftek M., Vincent C., Serre G.;
RT "Expression cloning of human corneodesmosin proves its identity with
RT the product of the S gene and allows improved characterization of its
RT processing during keratinocyte differentiation.";
RL J. Biol. Chem. 273:22640-22647(1998).
DR EMBL; AF030130; AAC24196.1; -.
DR PIR; A48679; A48679.
DR GO; GO:0005911; C:intercellular junction; TAS.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR GO; GO:0008544; P:epidermis development; TAS.
DR InterPro; IPR001917; Aminotrans_II.
DR PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
SQ SEQUENCE 529 AA; 51479 MW; 84418CFC465B584E CRC64;

Query Match 100.0%; Score 2752; DB 2; Length 529;
Best Local Similarity 100.0%; Pred. No. 1.3e-111;
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60

Qy         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180

Qy        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240

Qy        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
```

```
Qy        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360

Qy        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420

Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480

Qy        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 2
Q8IZU4
ID   Q8IZU4      PRELIMINARY;      PRT;   529 AA.
AC   Q8IZU4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Corneodesmosin.
GN   Name=CDSN;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22256218; PubMed=12366786;
RA   Hui J., Oka A., Tamiya G., Tomizawa M., Kulski J.K., Penhale W.J.,
RA   Tay G.K., Iizuka M., Ozawa A., Inoko H.;
RT   "Corneodesmosin DNA polymorphisms in MHC haplotypes and Japanese
RT   patients with psoriasis.";
RL   Tissue Antigens 60:77-83(2002).
DR   EMBL; AF491330; AAN70996.1; -.
DR   EMBL; AF491329; AAN70996.1; JOINED.
DR   InterPro; IPR001917; Aminotrans_II.
DR   PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
SQ   SEQUENCE   529 AA;  51495 MW;  844B8CEE465B585C CRC64;

  Query Match            99.9%;  Score 2749;  DB 2;  Length 529;
  Best Local Similarity  99.8%;  Pred. No. 1.7e-111;
  Matches  528;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60

Qy         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180

Qy        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240

Qy        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300

Qy        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360

Qy        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGSPYHPCGSASQ 420

Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480

Qy        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 3
Q7YR44
ID   Q7YR44      PRELIMINARY;      PRT;   529 AA.
AC   Q7YR44;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Modified desmosin.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22709134; PubMed=12799463; DOI=10.1073/pnas.1230533100;
RA   Anzai T., Shiina T., Kimura N., Yanagiya K., Kohara S., Shigenari A.,
RA   Yamagata T., Kulski J.K., Naruse T.K., Fujimori Y., Fukuzumi Y.,
RA   Yamazaki M., Tashiro H., Iwamoto C., Umehara Y., Imanishi T.,
RA   Meyer A., Ikeo K., Gojobori T., Bahram S., Inoko H.;
RT   "Comparative sequencing of human and chimpanzee MHC class I regions
RT   unveils insertions/deletions as the major path to genomic
RT   divergence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7708-7713(2003).
DR   EMBL; AB100083; BAC78170.1; -.
DR   InterPro; IPR001917; Aminotrans_II.
DR   PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
SQ   SEQUENCE   529 AA;  51611 MW;  7E7111EC28D0B018 CRC64;

  Query Match            99.2%;  Score 2731;  DB 2;  Length 529;
  Best Local Similarity  99.2%;  Pred. No. 1e-110;
  Matches  525;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPMRITSPNDPCLTGKG 60

Qy         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180

Qy        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NSYRGILNPSQPGQSSSSSQTFGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240

Qy        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300

Qy        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360

Qy        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGSLPYHPCGSASQ 420

Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
```

```
Db        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480

Qy        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 4
Q8IZU5
ID   Q8IZU5      PRELIMINARY;      PRT;   528 AA.
AC   Q8IZU5;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Corneodesmosin.
GN   Name=CDSN;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22256218; PubMed=12366786;
RA   Hui J., Oka A., Tamiya G., Tomizawa M., Kulski J.K., Penhale W.J.,
RA   Tay G.K., Iizuka M., Ozawa A., Inoko H.;
RT   "Corneodesmosin DNA polymorphisms in MHC haplotypes and Japanese
RT   patients with psoriasis.";
RL   Tissue Antigens 60:77-83(2002).
DR   EMBL; AF491328; AAN70995.1; -.
DR   EMBL; AF491327; AAN70995.1; JOINED.
DR   InterPro; IPR001917; Aminotrans_II.
DR   PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
SQ   SEQUENCE   528 AA;  51521 MW;  481ECE7D8CE07A38 CRC64;

  Query Match             98.8%;  Score 2719.5;  DB 2;  Length 528;
  Best Local Similarity   99.1%;  Pred. No. 3.2e-110;
  Matches  524;  Conservative    2;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60

Qy         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSG-SSSHSSSSSSFQFSSSSFQVGNGSALPTND 179

Qy        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        180 NSYRGILNPSQPGQSSSSSQTFGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 239

Qy        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 299

Qy        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 359

Qy        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||:| ||||||||
Db        360 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGLPYHPCGSASQ 419

Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
              |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        420 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSNSSGHPCMSVSSLTLTGGPDGSPHPDPS 479

Qy        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 528


RESULT 5
Q8IZU6
ID   Q8IZU6      PRELIMINARY;      PRT;   528 AA.
AC   Q8IZU6;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Corneodesmosin.
GN   Name=CDSN;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22256218; PubMed=12366786;
RA   Hui J., Oka A., Tamiya G., Tomizawa M., Kulski J.K., Penhale W.J.,
RA   Tay G.K., Iizuka M., Ozawa A., Inoko H.;
RT   "Corneodesmosin DNA polymorphisms in MHC haplotypes and Japanese
RT   patients with psoriasis.";
RL   Tissue Antigens 60:77-83(2002).
DR   EMBL; AF491326; AAN70994.1; -.
DR   EMBL; AF491325; AAN70994.1; JOINED.
DR   InterPro; IPR001917; Aminotrans_II.
DR   PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
SQ   SEQUENCE   528 AA;  51476 MW;  3C44FE7D96D35359 CRC64;

  Query Match             98.8%;  Score 2719.5;  DB 2;  Length 528;
  Best Local Similarity   99.1%;  Pred. No. 3.2e-110;
  Matches  524;  Conservative    2;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db          1 MGSSRAPWMGRVGGHGMLALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60

Qy         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSG-SSSHSSSSSSFQFSSSSFQVGNGSALPTND 179

Qy        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        180 NSYRGILNPSQPGQSSSSSQTFGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 239

Qy        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 299

Qy        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 359

Qy        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||:| ||||||||
Db        360 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGLPYHPCGSASQ 419

Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 479

Qy        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 528


RESULT 6
Q8N5P3
```

```
ID   Q8N5P3      PRELIMINARY;      PRT;   528 AA.
AC   Q8N5P3;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   CDSN protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RA   Strausberg R.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC031993; AAH31993.1; -.
DR   InterPro; IPR001917; Aminotrans_II.
DR   PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
SQ   SEQUENCE   528 AA;  51446 MW;  3C593E7456C75359 CRC64;

  Query Match             98.7%;  Score 2715.5;  DB 2;  Length 528;
  Best Local Similarity   98.9%;  Pred. No. 4.8e-110;
  Matches  523;  Conservative    2;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db          1 MGSSRAPWMGRVGGHGMLALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG 60

Qy         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSG 120

Qy        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTND 180
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        121 SGSSLQGASGSSQLGSSSSHSGSSGSHSG-SSSHSSSSSSFQFSSSSFQVGNGSALPTND 179

Qy        181 NSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 240
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        180 NSYRGILNPSQPGQSSSSSQTFGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYI 239

Qy        241 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 PSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVG 299

Qy        301 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 GSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQ 359

Qy        361 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQ 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db        360 SAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISGSSGLPYHPCGSASQ 419

Qy        421 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 SPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS 479

Qy        481 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 528


RESULT 7
Q95IF9
ID   Q95IF9      PRELIMINARY;      PRT;   513 AA.
AC   Q95IF9; Q86Z04;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   S protein (Modified desmosin expressed in non-cornified
DE   epidermis).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shiina S., Tamiya G., Oka A., Inoko H.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Hirakawa M., Yamaguchi H., Imai K., Shimada J.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Shiina T., Ota M., Katsuyama Y., Hashimoto N., Inoko H.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP000510; BAB63316.1; -.
DR   EMBL; AB088114; BAC54948.1; -.
DR   InterPro; IPR001917; Aminotrans_II.
DR   PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
SQ   SEQUENCE   513 AA;  49870 MW;  BFE1445FA533140A CRC64;

  Query Match             96.5%;  Score 2657;  DB 2;  Length 513;
  Best Local Similarity   99.8%;  Pred. No. 1.6e-107;
  Matches  512;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         17 MMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGS 76
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGS 60

Qy         77 SISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGS 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGS 120

Qy        137 SSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSS 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSS 180

Qy        197 SSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVV 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVV 240

Qy        257 VDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKG 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKG 300

Qy        317 KIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGG 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGG 360

Qy        377 VQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSS 436
              |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
```

```
Db        361 VQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGSPYHPCGSASQSPCSPPGTGSFSSSSS 420

Qy        437 SQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPC 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPC 480

Qy        497 RSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 529
              |||||||||||||||||||||||||||||||||
Db        481 RSIRDILAQVKPLGPQLADPEVFLPQGELLDSP 513


RESULT 8
Q9NP52
ID   Q9NP52      PRELIMINARY;      PRT;   500 AA.
AC   Q9NP52;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Corneodesmosin (Fragment).
GN   Name=CDSN;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21108612; PubMed=11169256;
RA   Guerrin M., Vincent C., Simon M., Ahnini R.T., Fort M., Serre G.;
RT   "Identification of six novel polymorphisms in the human corneodesmosin
RT   gene.";
RL   Tissue Antigens 57:32-38(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Guerrin-Weber M., Vincent C.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Guerrin-Weber M., Vincent C.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF224752; AAF78564.1; -.
DR   EMBL; AF224751; AAF78563.1; -.
DR   EMBL; AF286165; AAG02419.1; -.
DR   InterPro; IPR001917; Aminotrans_II.
DR   PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
FT   NON_TER       1       1
SQ   SEQUENCE   500 AA;  48575 MW;  F27C4A7761CBE966 CRC64;

  Query Match             94.3%;  Score 2596;  DB 2;  Length 500;
  Best Local Similarity   99.8%;  Pred. No. 6.6e-105;
  Matches  499;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         30 TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 89
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 60

Qy         90 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG 149
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG 120

Qy        150 SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQ 209
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQ 180

Qy        210 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQ 269
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQ 240

Qy        270 GPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENP 329
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENP 300

Qy        330 VKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKG 389
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKG 360

Qy        390 PCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCG 449
              |||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        361 PCSPSSSRVPSSSSISSSSGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCG 420

Qy        450 SKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL 509
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL 480

Qy        510 GPQLADPEVFLPQGELLDSP 529
              ||||||||||||||||||||
Db        481 GPQLADPEVFLPQGELLDSP 500


RESULT 9
Q9NPG5
ID   Q9NPG5      PRELIMINARY;      PRT;   500 AA.
AC   Q9NPG5; Q9UIN8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Corneodesmosin (Fragment).
GN   Name=CDSN;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21108612; PubMed=11169256;
RA   Guerrin M., Vincent C., Simon M., Ahnini R.T., Fort M., Serre G.;
RT   "Identification of six novel polymorphisms in the human corneodesmosin
RT   gene.";
RL   Tissue Antigens 57:32-38(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Guerrin-Weber M., Vincent C.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF224754; AAF78566.1; -.
DR   EMBL; AF224753; AAF78565.1; -.
DR   InterPro; IPR001917; Aminotrans_II.
DR   PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
FT   NON_TER       1       1
SQ   SEQUENCE   500 AA;  48619 MW;  43966B3A7F3E5278 CRC64;

  Query Match             94.2%;  Score 2593;  DB 2;  Length 500;
  Best Local Similarity   99.8%;  Pred. No. 8.9e-105;
  Matches  499;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         30 TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 89
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 60

Qy         90 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG 149
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG 120

Qy        150 SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQ 209
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTFGVSSSGQ 180

Qy        210 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQ 269
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQ 240

Qy        270 GPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENP 329
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENP 300

Qy        330 VKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKG 389
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKG 360
```

```
Qy        390 PCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCG 449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCG 420

Qy        450 SKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL 509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL 480

Qy        510 GPQLADPEVFLPQGELLDSP 529
              ||||||||||||||||||||
Db        481 GPQLADPEVFLPQGELLDSP 500


RESULT 10
Q9NRH8
ID   Q9NRH8      PRELIMINARY;      PRT;   500 AA.
AC   Q9NRH8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Corneodesmosin (Fragment).
GN   Name=CDSN;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21108612; PubMed=11169256;
RA   Guerrin M., Vincent C., Simon M., Ahnini R.T., Fort M., Serre G.;
RT   "Identification of six novel polymorphisms in the human corneodesmosin
RT   gene.";
RL   Tissue Antigens 57:32-38(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Guerrin-Weber M., Vincent C.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF224747; AAF78559.1; -.
DR   InterPro; IPR001917; Aminotrans_II.
DR   PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
FT   NON_TER       1       1
SQ   SEQUENCE   500 AA;  48635 MW;  439C0B287F3E526A CRC64;

  Query Match             94.1%;  Score 2590;  DB 2;  Length 500;
  Best Local Similarity   99.6%;  Pred. No. 1.2e-104;
  Matches  498;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         30 TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 60

Qy         90 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG 149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG 120

Qy        150 SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQ 209
              ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db        121 SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTFGVSSSGQ 180

Qy        210 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQ 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQ 240

Qy        270 GPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENP 329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENP 300

Qy        330 VKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKG 389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKG 360

Qy        390 PCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCG 449
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PCSPSSSRVPSSSSISSSSGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCG 420

Qy        450 SKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL 509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL 480

Qy        510 GPQLADPEVFLPQGELLDSP 529
              ||||||||||||||||||||
Db        481 GPQLADPEVFLPQGELLDSP 500


RESULT 11
Q9NRH6
ID   Q9NRH6      PRELIMINARY;      PRT;   500 AA.
AC   Q9NRH6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Corneodesmosin (Fragment).
GN   Name=CDSN;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21108612; PubMed=11169256;
RA   Guerrin M., Vincent C., Simon M., Ahnini R.T., Fort M., Serre G.;
RT   "Identification of six novel polymorphisms in the human corneodesmosin
RT   gene.";
RL   Tissue Antigens 57:32-38(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Guerrin-Weber M., Vincent C.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF224749; AAF78561.1; -.
DR   InterPro; IPR001917; Aminotrans_II.
DR   PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
FT   NON_TER       1       1
SQ   SEQUENCE   500 AA;  48669 MW;  DD06FE8D8FCBF79A CRC64;

  Query Match             94.0%;  Score 2586;  DB 2;  Length 500;
  Best Local Similarity   99.4%;  Pred. No. 1.8e-104;
  Matches  497;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         30 TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 89
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db          1 TLAKSIGTFSDPCKDPTRITSPNDPCFTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 60

Qy         90 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG 149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG 120

Qy        150 SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQ 209
              ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db        121 SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTFGVSSSGQ 180

Qy        210 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQ 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQ 240

Qy        270 GPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENP 329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENP 300

Qy        330 VKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKG 389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKG 360

Qy        390 PCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCG 449
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PCSPSSSRVPSSSSISSSSGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCG 420
```

```
Qy        450 SKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL 509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL 480

Qy        510 GPQLADPEVFLPQGELLDSP 529
              ||||||||||||||||||||
Db        481 GPQLADPEVFLPQGELLDSP 500


RESULT 12
Q9NRH7
ID   Q9NRH7      PRELIMINARY;      PRT;   500 AA.
AC   Q9NRH7; Q9UIN9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Corneodesmosin (Fragment).
GN   Name=CDSN;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21108612; PubMed=11169256;
RA   Guerrin M., Vincent C., Simon M., Ahnini R.T., Fort M., Serre G.;
RT   "Identification of six novel polymorphisms in the human corneodesmosin
RT   gene.";
RL   Tissue Antigens 57:32-38(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Guerrin-Weber M., Vincent C.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF224748; AAF78560.1; -.
DR   InterPro; IPR001917; Aminotrans_II.
DR   PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
FT   NON_TER       1      1
SQ   SEQUENCE   500 AA;  48605 MW;  4F1816E87F3E4FAA CRC64;

  Query Match             94.0%;  Score 2586;  DB 2;  Length 500;
  Best Local Similarity   99.4%;  Pred. No. 1.8e-104;
  Matches  497;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         30 TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 60

Qy         90 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG 149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG 120

Qy        150 SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQ 209
              ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db        121 SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTFGVSSSGQ 180

Qy        210 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQ 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQ 240

Qy        270 GPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENP 329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENP 300

Qy        330 VKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKG 389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKG 360

Qy        390 PCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCG 449
              |||||||||| ||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        361 PCSPSSSRVPSGSSISSSSGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCG 420

Qy        450 SKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL 509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL 480

Qy        510 GPQLADPEVFLPQGELLDSP 529
              ||||||||||||||||||||
Db        481 GPQLADPEVFLPQGELLDSP 500


RESULT 13
Q9NPE0
ID   Q9NPE0      PRELIMINARY;      PRT;   500 AA.
AC   Q9NPE0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Corneodesmosin (Fragment).
GN   Name=CDSN;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21108612; PubMed=11169256;
RA   Guerrin M., Vincent C., Simon M., Ahnini R.T., Fort M., Serre G.;
RT   "Identification of six novel polymorphisms in the human corneodesmosin
RT   gene.";
RL   Tissue Antigens 57:32-38(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Guerrin-Weber M., Vincent C.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF224758; AAF78570.1; -.
DR   EMBL; AF224756; AAF78568.1; -.
DR   EMBL; AF224757; AAF78569.1; -.
DR   InterPro; IPR001917; Aminotrans_II.
DR   PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
FT   NON_TER       1      1
SQ   SEQUENCE   500 AA;  48687 MW;  69BCF6FB7C1DA265 CRC64;

  Query Match             93.6%;  Score 2576;  DB 2;  Length 500;
  Best Local Similarity   99.0%;  Pred. No. 4.8e-104;
  Matches  495;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         30 TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 60

Qy         90 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG 149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG 120

Qy        150 SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQ 209
              :|||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db        121 NSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTFGVSSSGQ 180

Qy        210 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQ 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQ 240

Qy        270 GPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENP 329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENP 300

Qy        330 VKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKG 389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKG 360

Qy        390 PCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCG 449
              |||||||||||||||:| ||||||||||||||||||||||||||||||||||||||||||
Db        361 PCSPSSSRVPSSSSISSSSGLPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCG 420

Qy        450 SKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL 509
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL 480

Qy        510 GPQLADPEVFLPQGELLDSP 529
                ||||||||||||||||||:||
Db        481 GPQLADPEVFLPQGELLNSP 500


RESULT 14
Q9NRH5
ID   Q9NRH5      PRELIMINARY;      PRT;   499 AA.
AC   Q9NRH5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Corneodesmosin (Fragment).
GN   Name=CDSN;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21108612; PubMed=11169256;
RA   Guerrin M., Vincent C., Simon M., Ahnini R.T., Fort M., Serre G.;
RT   "Identification of six novel polymorphisms in the human corneodesmosin
RT   gene.";
RL   Tissue Antigens 57:32-38(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Guerrin-Weber M., Vincent C.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF224750; AAF78562.1; -.
DR   InterPro; IPR001917; Aminotrans_II.
DR   PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
FT   NON_TER       1       1
SQ   SEQUENCE   499 AA;  48570 MW;  BCF53C0C65FE2EB8 CRC64;

  Query Match             93.4%;  Score 2571.5;  DB 2;  Length 499;
  Best Local Similarity   99.2%;  Pred. No. 7.5e-104;
  Matches  496;  Conservative    1;  Mismatches    2;  Indels    1;  Gaps    1;

Qy         30 TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 89
              ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db          1 TLAKSIGTFSDPCKDPTRITSPNDPCFTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 60

Qy         90 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG 149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG 120

Qy        150 SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQ 209
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        121 SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTFGVSSSGQ 180

Qy        210 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQ 269
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        181 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRP-VVVVDQHGSGAPGVVQ 239

Qy        270 GPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENP 329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 GPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENP 299

Qy        330 VKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKG 389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 VKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKG 359

Qy        390 PCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCG 449
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        360 PCSPSSSRVPSSSSISSSSGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCG 419

Qy        450 SKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL 509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 SKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL 479

Qy        510 GPQLADPEVFLPQGELLDSP 529
              ||||||||||||||||||||
Db        480 GPQLADPEVFLPQGELLDSP 499


RESULT 15
Q9NRH4
ID   Q9NRH4      PRELIMINARY;      PRT;   499 AA.
AC   Q9NRH4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Corneodesmosin (Fragment).
GN   Name=CDSN;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21108612; PubMed=11169256;
RA   Guerrin M., Vincent C., Simon M., Ahnini R.T., Fort M., Serre G.;
RT   "Identification of six novel polymorphisms in the human corneodesmosin
RT   gene.";
RL   Tissue Antigens 57:32-38(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Guerrin-Weber M., Vincent C.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF224755; AAF78567.1; -.
DR   InterPro; IPR001917; Aminotrans_II.
DR   PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
FT   NON_TER       1       1
SQ   SEQUENCE   499 AA;  48574 MW;  20E67AC82D81C013 CRC64;

  Query Match             93.4%;  Score 2569.5;  DB 2;  Length 499;
  Best Local Similarity   99.2%;  Pred. No. 9.2e-104;
  Matches  496;  Conservative    1;  Mismatches    2;  Indels    1;  Gaps    1;

Qy         30 TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSS 60

Qy         90 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG 149
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db         61 GSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSG-SGSHSG 119

Qy        150 SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQ 209
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        120 SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTFGVSSSGQ 179

Qy        210 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQ 269
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        180 SVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQ 239

Qy        270 GPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENP 329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 GPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENP 299

Qy        330 VKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKG 389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 VKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKG 359

Qy        390 PCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCG 449
              ||||||||||||||||||:| ||||||||||||||||||||||||||||||||||||||||
Db        360 PCSPSSSRVPSSSSISSSSGLPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCG 419

Qy        450 SKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL 509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 SKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL 479

Qy        510 GPQLADPEVFLPQGELLDSP 529
```

```
                    ||||||||||||||||||||
Db        480 GPQLADPEVFLPQGELLDSP 499

Search completed: August  2, 2005, 09:12:21
Job time : 177 secs
```